The Dissertation Committee for Saumik P. Dana
certifies that this is the approved version of the following dissertation:

Addressing challenges in modeling of coupled flow and poromechanics in deep subsurface reservoirs

Committee:

Mary F. Wheeler, Supervisor

Chad Landis

Rui Huang

Matthew Balhoff

Addressing challenges in modeling of coupled flow and poromechanics in deep subsurface reservoirs

by

Saumik P. Dana

DISSERTATION

Presented to the Faculty of the Graduate School of
The University of Texas at Austin
in Partial Fulfillment
of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN
December 2018

Acknowledgments

I would like to express my gratitude to Prof. Mary F. Wheeler for providing me with the opportunity to work on my doctoral dissertation under her supervision. I would like to thank Dr. Benjamin Ganis for the help he provided in getting the parallel implementation going in IPARS, which is quite a tough nut to crack in itself. I would like to thank Dr. Gurpreet Singh for the mentorship he provided towards the latter half of my tenure at the Center for Subsurface Modeling. I would like to thank Prof. Kundan Kumar for the couple of discussions that sparked off the theoretical works following the first journal publication. I would like to thank Dr. Tameem Almani for the initial moral support when I was trying to get my bearings in the research group. Thanks to the folks I met at UT: Nishant, Darshan, Chinmaya, Nagaraja, Sparsh, Roy, Kulkarni, Agarwal, Gupta, Barik, Subramanian, Krishnamurty, Lo, Soham, Shubham, Verma, Kakkar, Anand, Sundeep, Prem, Sharma, Vivek, Pushkar, Sriram, Abhimanyu, Andrei, and Martin. Finally, I would like to thank my Mom for being party to my rants and monologues through my six year stay at UT.

Addressing challenges in modeling of coupled flow and poromechanics in deep subsurface reservoirs

Publication No. ____________________

Saumik P. Dana, Ph.D.
The University of Texas at Austin, 2018

Supervisor: Mary F. Wheeler

In coupled flow and poromechanics phenomena representing hydrocarbon production or CO_2 sequestration in deep subsurface *non-fractured* reservoirs, the spatial domain in which fluid flow occurs is usually much smaller than the spatial domain over which significant deformation occurs. The vertical extent of the poromechanical domain can be two orders of magnitude more than the characteristic thickness of the flow domain (reservoir). The lateral extent of the poromechanical domain should also be allowed to be substantially larger than that of the flow domain to enable the imposition of far-field boundary conditions on the poromechanical domain. The typical approach is to either impose an overburden pressure directly on the reservoir thus treating it as a coupled problem domain or to model flow on a huge domain with zero permeability cells to mimic the no flow boundary condition on the interface of the reservoir and the surrounding rock. The former approach precludes a study of land subsidence or uplift and further does not mimic the true effect of the overburden on stress sensitive reservoirs whereas the latter approach has huge computational costs. The flow domain requires an areal resolution fine enough to be able to capture the underlying nonlinearities in the multiphase flow equations. If the same grid resolution is employed for the poromechanical domain, the simulator would crash for lack of memory and computing resource. With that in mind, it is imperative to establish a framework in which fluid flow is resolved on a finer grid and poromechanical deformation is resolved on a coarse grid. In addition, the geometry of the flow domain necessitates the use of non-nested grids which allows for freedom of choice of the poromechanical grid resolution. Furthermore, to achieve the goal of rendering realistic simulations of subsurface phenomena, we cannot ignore the heterogeneity in flow and poromechanical properties, as well as the lack in accuracy of the poromechanical calculations if the grid for the poromechanics domain is too coarse. This dissertation is a rendition of how we invoke concepts in computational geometry, parallel computing, applied mathematics and convex optimization in designing and implementing algorithms that tackle all the aforementioned challenges.

Table of Contents

Chapter 3. Construction of the up and downscaling operators and verification of the two-grid algorithm 33

Chapter 4. Convergence analysis of the two-grid algorithm and the link with computational homogenization for heterogeneous poroelastic media 66

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Problem statement for the dissertation

The problem statement for this dissertation is to enable the study of the interaction between fluid flow in deep subsurface reservoirs and poromechanical (or geomechanical) deformation using a parallel finite element framework as shown in Figure 1.1. The tasks we perform are:

- Use the ideas of MPI based parallel computing to decompose the problem domain into a flow domain and a poromechanical domain with the flow domain being a subset of the poromechanical domain. This general approach would also work for the specific case of the two domains being identical.
- Construct mapping operators that enable the implementation of a two-grid approach in which the flow equations are solved on a fine mesh and the poromechanical equations are solved on a coarse mesh. Modify the single-grid solution algorithm that seamlessly incorporates the two-grid approach
- Demonstrate numerically and theoretically the convergence of the modified algorithm for consolidation in homogeneous, as well as heterogeneous

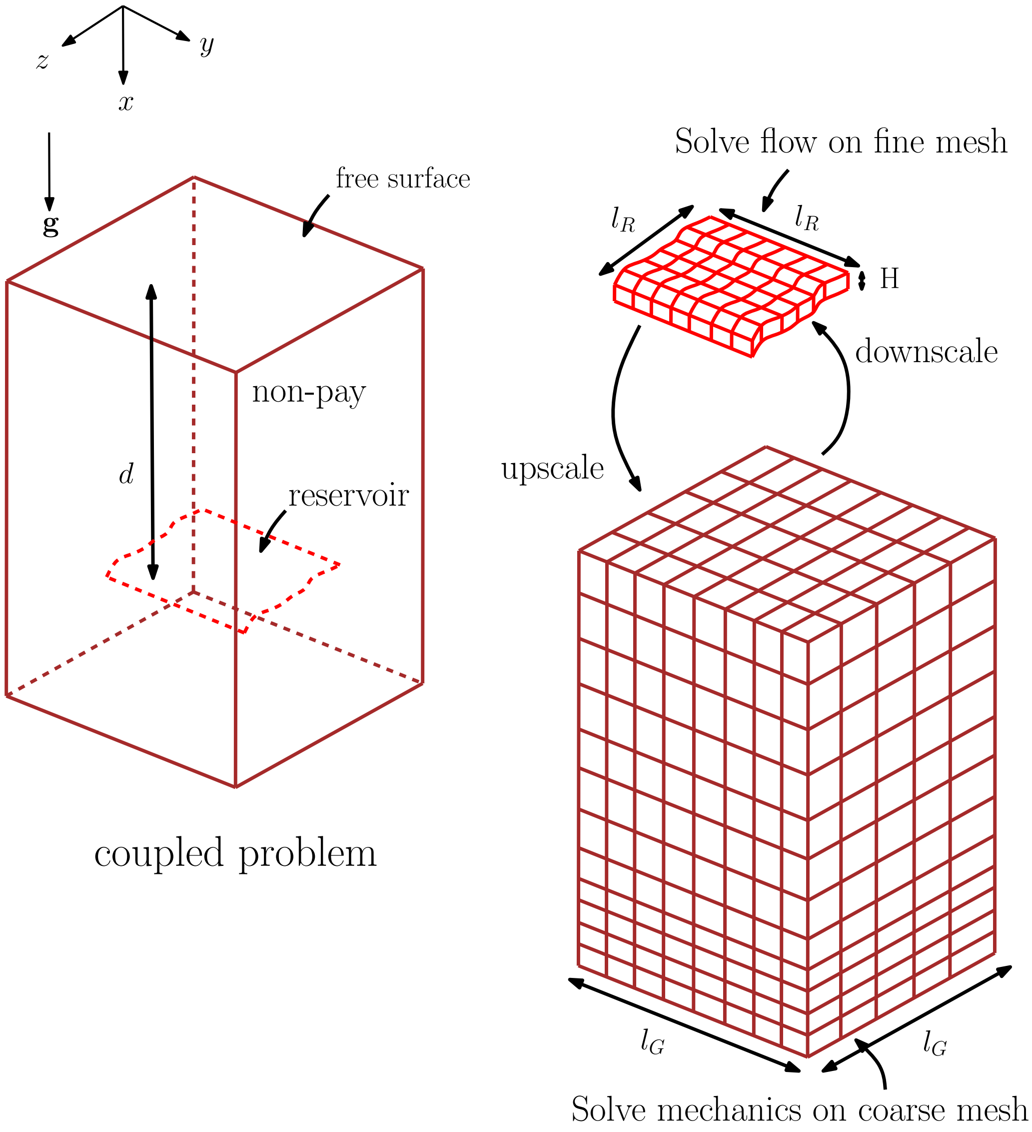

Figure 1.1: l_R and l_G are the characteristic lateral extents whereas d and H are the characteristic vertical extents of the flow and poromechanical domains respectively. Typically, both l_R and d are in the range of 1000 ft to 10000 ft and $H \sim 100\, ft$.

porous media, for the case of isotropic poroelasticity

In addition to that, we also perform the following tasks

- Study the convergence properties of the standard single-grid approach for the case of the anisotropic poroelasticity. Eventually, in subsurface problems, it is imperative to factor in the anisotropy of the underlying rocks
- Formulate an optimization method for treatment of hanging nodes in the mechanics mesh. These hanging nodes would allow for aggressive coarsening inside the mechanics mesh around the region surrounding the reservoir with the coarse mesh extending all the way to the surface

1.2 Outline of the dissertation

In Chapter 2, we present the model equations and discrete variational statements for the two-grid staggered solution algorithm. Since flow and poromechanics are solved on different finite element grids, we design grid-to-grid projection operators using concepts in computational geometry and parallel computing. The design of these operators is the subject of Chapter 3. In Chapter 4, we invoke the concept of contraction mapping to establish the theoretical convergence of the solution algorithm presented in Chapter 3 and further generalize the algorithm in the presence of heterogeneities in the porous medium. The work in Chapters 2, 3, and 4 assumes isotropy of the porous solid. In Chapter 5, we again invoke the concept of contraction mapping to establish the theoretical convergence of the staggered solution algorithm with the anisotropy of the porous rock factored in. In Chapter 6, we present an augmented Lagrangian formulation for treatment of hanging nodes in hexahedral meshes. In Chapter 7, we present conclusions and scope for future work.

1.3 Journal articles resulting from the dissertation

The following articles were spawned out of the body of work comprising this dissertation

- S. Dana, J. Ita and M. F. Wheeler, "The correspondence between Voigt and Reuss bonds and the decoupling constraint in a two-grid staggered solution algorithm for coupled flow and deformation in heterogeneous poroelastic media", *arXiv* 1810.09443
- S. Dana and M. F. Wheeler, "Augmented Lagrangian for treatment of hanging nodes in hexahedral meshes", *arXiv* 1809.04031
- S. Dana and M. F. Wheeler, "Convergence analysis of fixed stress split iterative scheme for anisotropic poroelasticity with tensor Biot parameter", *Computational Geosciences*, 22(5), 1219-1230, 2018
- S. Dana and M. F. Wheeler, "Convergence analysis of two-grid fixed stress split iterative scheme for coupled flow and deformation in heterogeneous poroelastic media", *Computer Methods in Applied Mechanics and Engineering*, 341:788-806, 2018

- S. Dana, B. Ganis and M. F. Wheeler, “A multiscale fixed stress split iterative scheme for coupled flow and poromechanics in deep subsurface reservoirs”, *Journal of Computational Physics*, 352:1-22, 2018

Chapter 2

Model equations and discrete variational statements for the two-grid staggered solution algorithm

2.1 Model equations

The equations for slightly single compressible single phase flow coupled with linear poromechanics are given in this module

2.1.1 Flow model

Let $\Omega^f \subset \mathbb{R}^3$ be the flow domain with boundary $\partial\Omega^f = \Gamma_D^f \cup \Gamma_N^f$ where Γ_D^f is Dirichlet boundary and Γ_N^f is Neumann boundary. The mass conservation equation (2.1) for single phase flow in porous medium[1] with the Darcy law (2.2) for slightly compressible fluid (2.3) with boundary conditions (2.4) and initial conditions (2.5) is

$$\frac{\partial(\phi^*\rho)}{\partial t} + \nabla \cdot \mathbf{z} = q \tag{2.1}$$

$$\mathbf{z} = -\frac{\mathbf{K}\rho}{\mu}(\nabla p - \rho\mathbf{g}) \tag{2.2}$$

$$\rho = \rho_0 e^{c\,(p-p_0)} \tag{2.3}$$

$$p = g \text{ on } \Gamma_D^f \times (0,T], \;\; \mathbf{z}\cdot\mathbf{n} = 0 \text{ on } \Gamma_N^f \times (0,T] \tag{2.4}$$

---

[1] The derivation of the equation is given in Coussy [24]

$$p(\mathbf{x},0) = p_0(\mathbf{x}),\ \rho(\mathbf{x},0) = \rho_0(\mathbf{x}),\ \phi(\mathbf{x},0) = \phi_0(\mathbf{x}) \qquad \forall \mathbf{x} \in \Omega^f \tag{2.5}$$

where $p : \Omega^f \times (0,T] \to \mathbb{R}$ is the fluid pressure, $\mathbf{z} : \Omega^f \times (0,T] \to \mathbb{R}^3$ is the fluid flux, $\bar{\epsilon}$ is the volumetric strain, ϕ is the Eulerian porosity defined as the ratio of the pore volume in the deformed configuration to the total volume in the deformed configuration, $\phi^* \equiv \phi(1+\bar{\epsilon})$ is the Lagrangian porosity[2] defined as the ratio of the pore volume in the deformed configuration to the total volume in the undeformed configuration, $\mathbf{n}$ is the unit outward normal on Γ_N^f, q is the source or sink term, $\mathbf{K}$ is the uniformly symmetric positive definite absolute permeability tensor, μ is the fluid viscosity, ρ_0 is a reference density, c is the fluid compressibility and $T > 0$ is the time interval.

2.1.2 Poromechanics model

Let $\Omega^p \subset \mathbb{R}^3$ be the poromechanics domain with boundary $\partial\Omega^p = \Gamma_D^p \cup \Gamma_N^p$ where Γ_D^p is Dirichlet boundary and Γ_N^p is Neumann boundary. The linear momentum balance in the quasi-static limit of interest (2.6) with the definition of the total stress (2.7) (see Biot [8]) with the expression for the body force (2.8) and the small strain assumption (2.9) with boundary conditions (2.10) and initial condition (2.11) is

$$\nabla \cdot \boldsymbol{\sigma} + \mathbf{f} = \mathbf{0} \tag{2.6}$$

$$\boldsymbol{\sigma} = \mathbb{D}\boldsymbol{\epsilon} - \alpha p \mathbf{I} \equiv \lambda \bar{\epsilon} \mathbf{I} + 2G\boldsymbol{\epsilon} - \alpha p \mathbf{I} \tag{2.7}$$

---

[2] The reader is again refered to Coussy [24] for a more detailed explanation of Eulerian and Lagrangian porosities

$$\mathbf{f} = \rho\phi\mathbf{g} + \rho_r(1-\phi)\mathbf{g} \tag{2.8}$$

$$\boldsymbol{\epsilon}(\mathbf{u}) = \frac{1}{2}(\nabla\mathbf{u} + \nabla^T\mathbf{u}) \tag{2.9}$$

$$\mathbf{u}\cdot\mathbf{n}_1 = 0 \text{ on } \Gamma_D^p \times [0,T],\ \boldsymbol{\sigma}^T\mathbf{n}_2 = \mathbf{t} \text{ on } \Gamma_N^p \times [0,T] \tag{2.10}$$

$$\mathbf{u}(\mathbf{x},0) = \mathbf{0} \qquad \forall\ \mathbf{x} \in \Omega^p \tag{2.11}$$

where $\mathbf{u} : \Omega^p \times [0,T] \to \mathbb{R}^3$ is the solid displacement, ρ_r is the rock density, G is the shear modulus, λ is the Lame parameter, $\mathbf{n}_1$ is the unit outward normal to Γ_D^p, $\mathbf{n}_2$ is the unit outward normal to Γ_N^p, α is the Biot parameter,$\mathbf{f}$ is body force per unit volume, $\mathbf{t}$ is the traction boundary condition, $\boldsymbol{\epsilon}$ is the strain tensor, $\mathbb{D}$ is the fourth order elasticity tensor and $\mathbf{I}$ is the second order identity tensor.

2.1.3 Expression for mean stress

The mean stress σ_v is given as

$$\begin{aligned}\sigma_v &= \frac{1}{3}tr(\boldsymbol{\sigma}) = \frac{1}{3}tr\big(\lambda\bar{\epsilon}\mathbf{I} + 2G\boldsymbol{\epsilon} - \alpha p\mathbf{I}\big) = \frac{1}{3}3\lambda\bar{\epsilon} + \frac{1}{3}2G\,\overbrace{tr(\boldsymbol{\epsilon})}^{\bar{\epsilon}} - \frac{1}{3}3\alpha p \\ &= \big(\lambda + \frac{1}{3}2G\big)\bar{\epsilon} - \alpha p = K_b\bar{\epsilon} - \alpha p \end{aligned} \tag{2.12}$$

where $K_b \equiv \lambda + \frac{1}{3}2G$ is the bulk modulus of the solid skeleton or the drained bulk modulus.

2.2 The fixed stress split staggered solution strategy

Staggered solution schemes are those in which an operator splitting strategy is used to split the coupled problem into well-posed flow and mechanics subproblems which are then solved sequentially instead of solving the coupled system of equations monolithically. The fixed stress split is one such strategy that solves the flow problem while freezing the mean stress followed by the mechanics problem. This procedure (also refered to as ‘coupling iteration’ or ‘fixed - stress iteration’) is repeated until a certain convergence criterion is met. The basic idea of the fixed - stress split strategy is to solve for the pressure and flux degrees of freedom at the current fixed - stress iteration based on the value of the mean stress from the previous fixed - stress iteration. These pressures then contribute to the force vector in the poromechanics system which is solved for displacements thereby updating the stress state. This updated stress state is then fed back to the flow system for the next fixed - stress iteration. Since this strategy condemns the porous solid to follow a certain stress path during the flow solve, the convergence of the solution algorithm is not automatically guaranteed. Mikelić and Wheeler [58] proved the convergence of the fixed - stress split strategy using the principle of contraction mapping with appropriately chosen metrics. Castelletto et al. [20] showed that the fixed stress split strategy can be interpreted as a block triangular preconditioning strategy applied within a Richardson iteration to the fully implicit scheme. They also showed that a mixed formulation for flow with the fixed mean stress constraint coupled with a continuous Galerkin formulation for poromechanics is first order accurate in space and time for both pressure and displacement, even for distorted meshes.

2.3 Discrete variational statements

Given a domain $\Omega \subset \mathbb{R}^n$, we use $\mathbb{P}_k(\Omega)$ to represent the restriction of the space of polynomials of degree less that or equal to k to Ω. Given a domain $\Omega \subset \mathbb{R}^3$, we use $\mathbb{Q}_1(\Omega)$ to denote the space of trilinears[3] on Ω. The Sobolev spaces (Adams [2]) are based on the space of square integrable functions on a domain $\Omega \subset \mathbb{R}^3$ given by

$$L^2(\Omega) \equiv \left\{ f : \int\limits_{\Omega} |f|^2 = \|f\|^2_{L^2(\Omega)} < +\infty \right\},$$

We then define in general, for any integer $m \geq 0$

$$H^m(\Omega) \equiv \left\{ w : D^\alpha w \in L^2(\Omega) \ \forall |\alpha| \leq m \right\},$$

where the derivatives are taken in the sense of distributions and given by

$$D^\alpha w = \frac{\partial^{|\alpha|} w}{\partial x_1^{\alpha_1} .. \partial x_n^{\alpha_n}}, \ |\alpha| = \alpha_1 + \cdots + \alpha_n,$$

In the presence of a divergence operator in second order PDEs, the following space is defined

$$\mathbf{H}(div, \Omega) \equiv \left\{ \mathbf{v} : \mathbf{v} \in (L^2(\Omega))^3, \nabla \cdot \mathbf{v} \in L^2(\Omega) \right\}$$

For $\Omega \subset \mathbb{R}^3$ and $(\cdot)$ defined on Ω, we omit the differential dV in the representation of the integral of $(\cdot)$ over Ω for the sake of brevity as follows

$$\int\limits_{\Omega} (\cdot) \equiv \int\limits_{\Omega} (\cdot) \, dV$$

---

[3] If $\Omega \subset \mathbb{R}^2$, $\mathbb{Q}_1(\Omega)$ is used to denote the space of bilinears on Ω.

For $\Omega \subset \mathbb{R}^2$ and $(\cdot)$ defined on Ω, we omit the differential dS in the representation of the integral of $(\cdot)$ over Ω for the sake of brevity as follows

$$\int_{\Omega} (\cdot) \equiv \int_{\Omega} (\cdot)\, dS$$

2.3.1 Finite element mapping

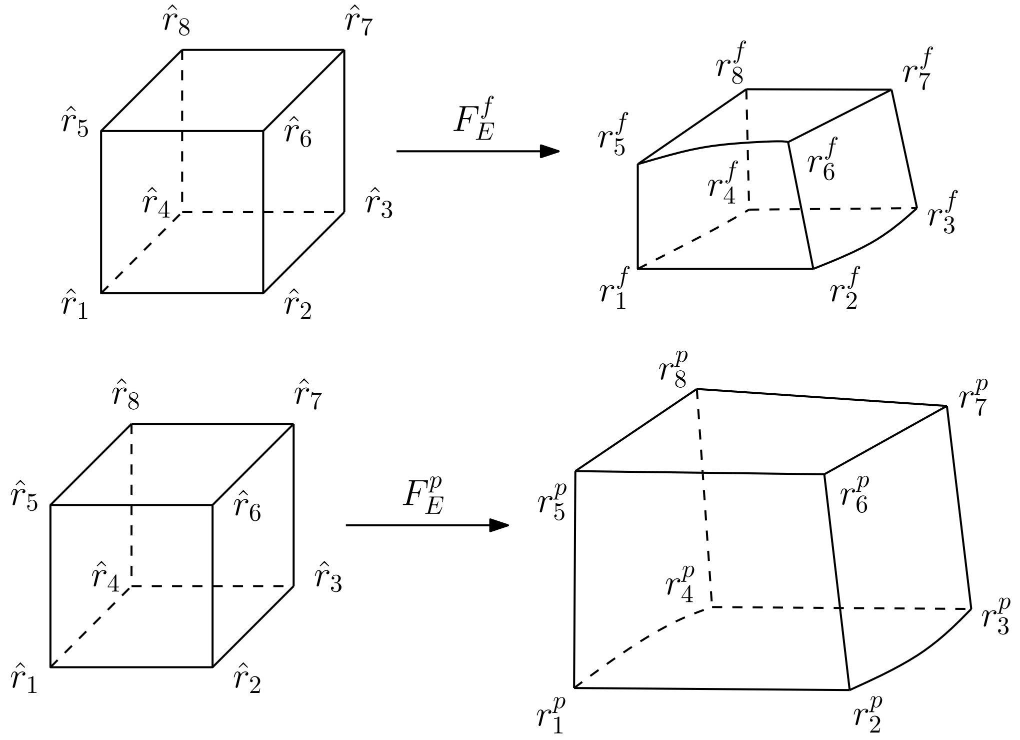

Figure 2.1: Trilinear mapping $F_E^f : \hat{E} \mapsto E^f$ and $F_E^p : \hat{E} \mapsto E^p$ for 8 noded distorted hexahedral elements E^f and E^p respectively. The faces of E^f and E^p can be non-planar.

Let $\mathscr{T}_h^f$ be the finite element partition of Ω^f consisting of distorted hexahedral elements E^f and let $\mathscr{T}_H^p$ be the finite element partition of Ω^p consisting of distorted hexahedral elements E^p. As the flow model is solved on a finer mesh, the mesh ratio r is such that

$$r \equiv \max_{E^p \in \mathscr{T}_H^p} diam(E^p) / \max_{E^f \in \mathscr{T}_h^f} diam(E^f) = H/h \geq 1$$

Let $\mathbf{r}_i^f$, $i = 1, .., 8$ and $\mathbf{r}_i^p$, $i = 1, .., 8$ be the vertices of E^f and E^p respectively. Let $\hat{E}$ represent a reference unit cube with vertices $\hat{\mathbf{r}}_1 \equiv (0,0,0)$, $\hat{\mathbf{r}}_2 \equiv (1,0,0)$, $\hat{\mathbf{r}}_3 \equiv (1,1,0)$, $\hat{\mathbf{r}}_4 \equiv (0,1,0)$, $\hat{\mathbf{r}}_5 \equiv (0,0,1)$, $\hat{\mathbf{r}}_6 \equiv (1,0,1)$, $\hat{\mathbf{r}}_7 \equiv (1,1,1)$ and $\hat{\mathbf{r}}_8 \equiv (0,1,1)$. Letting $\hat{\mathbf{x}} = (\hat{x}, \hat{y}, \hat{z}) \in \hat{E}$, the functions $F_E^f(\hat{\mathbf{x}}) : \hat{E} \mapsto E^f$ and $F_E^p(\hat{\mathbf{x}}) : \hat{E} \mapsto E^p$ as shown in Figure 2.1 are defined as

$$\begin{aligned} F_E^f(\hat{\mathbf{x}}) = {} & \mathbf{r}_1^f(1-\hat{x})(1-\hat{y})(1-\hat{z}) + \mathbf{r}_2^f\hat{x}(1-\hat{y})(1-\hat{z}) + \mathbf{r}_3^f\hat{x}\hat{y}(1-\hat{z}) \\ & + \mathbf{r}_4^f(1-\hat{x})\hat{y}(1-\hat{z}) + \mathbf{r}_5^f(1-\hat{x})(1-\hat{y})\hat{z} + \mathbf{r}_6^f\hat{x}(1-\hat{y})\hat{z} \\ & + \mathbf{r}_7^f\hat{x}\hat{y}\hat{z} + \mathbf{r}_8^f(1-\hat{x})\hat{y}\hat{z} \end{aligned} \tag{2.13}$$

$$\begin{aligned} F_E^p(\hat{\mathbf{x}}) = {} & \mathbf{r}_1^p(1-\hat{x})(1-\hat{y})(1-\hat{z}) + \mathbf{r}_2^p\hat{x}(1-\hat{y})(1-\hat{z}) + \mathbf{r}_3^p\hat{x}\hat{y}(1-\hat{z}) \\ & + \mathbf{r}_4^p(1-\hat{x})\hat{y}(1-\hat{z}) + \mathbf{r}_5^p(1-\hat{x})(1-\hat{y})\hat{z} + \mathbf{r}_6^p\hat{x}(1-\hat{y})\hat{z} \\ & + \mathbf{r}_7^p\hat{x}\hat{y}\hat{z} + \mathbf{r}_8^p(1-\hat{x})\hat{y}\hat{z} \end{aligned} \tag{2.14}$$

We denote the jacobian of F_E^f and F_E^p by DF_E^f and DF_E^p respectively and the determinant of DF_E^f and DF_E^p by J_E^f and J_E^p respectively. Defining $\mathbf{r}_{ij}^f \equiv \mathbf{r}_i^f - \mathbf{r}_j^f$ and $\mathbf{r}_{ij}^p \equiv \mathbf{r}_i^p - \mathbf{r}_j^p$, we have

$$DF_E^f(\hat{\mathbf{x}}) = \begin{bmatrix} \mathbf{r}_{21}^f + (\mathbf{r}_{34}^f - \mathbf{r}_{21}^f)\hat{y} + (\mathbf{r}_{65}^f - \mathbf{r}_{21}^f)\hat{z} + ((\mathbf{r}_{21}^f - \mathbf{r}_{34}^f) - (\mathbf{r}_{65}^f - \mathbf{r}_{78}^f))\hat{y}\hat{z}; \\ \mathbf{r}_{41}^f + (\mathbf{r}_{34}^f - \mathbf{r}_{21}^f)\hat{x} + (\mathbf{r}_{85}^f - \mathbf{r}_{41}^f)\hat{z} + ((\mathbf{r}_{21}^f - \mathbf{r}_{34}^f) - (\mathbf{r}_{65}^f - \mathbf{r}_{78}^f))\hat{x}\hat{z}; \\ \mathbf{r}_{51}^f + (\mathbf{r}_{65}^f - \mathbf{r}_{21}^f)\hat{x} + (\mathbf{r}_{85}^f - \mathbf{r}_{41}^f)\hat{y} + ((\mathbf{r}_{21}^f - \mathbf{r}_{34}^f) - (\mathbf{r}_{65}^f - \mathbf{r}_{78}^f))\hat{x}\hat{y} \end{bmatrix}_{3\times 3}$$

$$DF_E^p(\hat{\mathbf{x}}) = \begin{bmatrix} \mathbf{r}_{21}^p + (\mathbf{r}_{34}^p - \mathbf{r}_{21}^p)\hat{y} + (\mathbf{r}_{65}^p - \mathbf{r}_{21}^p)\hat{z} + ((\mathbf{r}_{21}^p - \mathbf{r}_{34}^p) - (\mathbf{r}_{65}^p - \mathbf{r}_{78}^p))\hat{y}\hat{z}; \\ \mathbf{r}_{41}^p + (\mathbf{r}_{34}^p - \mathbf{r}_{21}^p)\hat{x} + (\mathbf{r}_{85}^p - \mathbf{r}_{41}^p)\hat{z} + ((\mathbf{r}_{21}^p - \mathbf{r}_{34}^p) - (\mathbf{r}_{65}^p - \mathbf{r}_{78}^p))\hat{x}\hat{z}; \\ \mathbf{r}_{51}^p + (\mathbf{r}_{65}^p - \mathbf{r}_{21}^p)\hat{x} + (\mathbf{r}_{85}^p - \mathbf{r}_{41}^p)\hat{y} + ((\mathbf{r}_{21}^p - \mathbf{r}_{34}^p) - (\mathbf{r}_{65}^p - \mathbf{r}_{78}^p))\hat{x}\hat{y} \end{bmatrix}_{3\times 3}$$

Let $\mathbf{x}^f \in E^f$ and $\mathbf{x}^p \in E^p$ and further let $\psi(\mathbf{x}^f)$ and $\psi(\mathbf{x}^p)$ be functions defined on E^f and E^p respectively. Let $\hat{\psi}(\hat{\mathbf{x}})$ be the corresponding definitions on $\hat{E}$. Then we have

$$\nabla\psi(\mathbf{x}^f) = (DF_E^f)^{-T}\,\hat{\nabla}\hat{\psi} \tag{2.15}$$

$$\nabla\psi(\mathbf{x}^p) = (DF_E^p)^{-T}\,\hat{\nabla}\hat{\psi} \tag{2.16}$$

2.3.2 Mixed finite element spaces employed

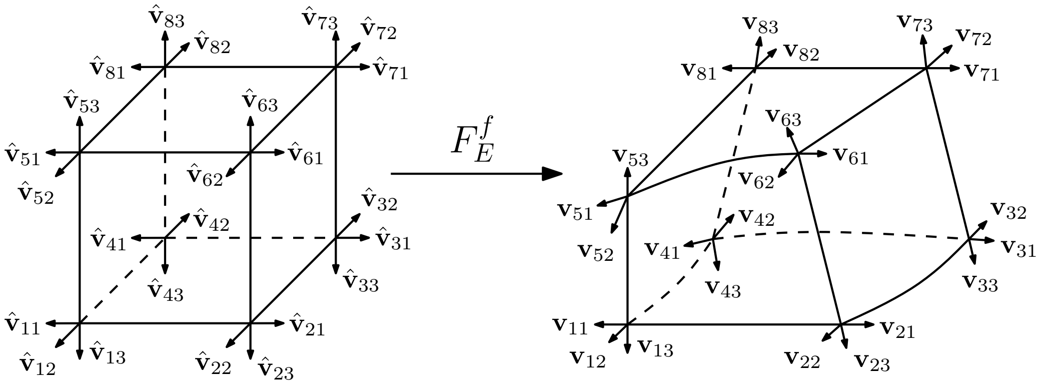

Figure 2.2: Degrees of freedom and basis functions for the enhanced BDDF$_1$ velocity space on hexahedra.

Let $\mathbf{V}_h^* \times W_h$ be the lowest order BDDF$_1$ MFE spaces on hexahedra (see [16]). With $\mathbf{x} \equiv (\hat{x}, \hat{y}, \hat{z}) \in \hat{E}$, these spaces are defined on $\hat{E}$ as

$$\begin{aligned}\hat{\mathbf{V}}^*(\hat{E}) &= (\mathbb{P}_1(\hat{E}))^3 + r_0\, curl(0,0,\hat{x}\hat{y}\hat{z})^T + r_1\, curl(0,0,\hat{x}\hat{y}^2)^T + s_0\, curl(\hat{x}\hat{y}\hat{z},0,0)^T \\ &\quad + s_1\, curl(\hat{y}\hat{z}^2,0,0)^T + t_0\, curl(0,\hat{x}\hat{y}\hat{z},0)^T + t_1\, curl(0,\hat{x}^2\hat{z},0)^T \\ \hat{W}(\hat{E}) &= \mathbb{P}_0(\hat{E})\end{aligned}$$

with the following properties

$$\hat{\nabla} \cdot \hat{\mathbf{V}}^*(\hat{E}) = \hat{W}(\hat{E}), \qquad \text{and} \qquad \forall \hat{\mathbf{v}} \in \hat{\mathbf{V}}^*(\hat{E}),\ \forall \hat{e} \subset \partial \hat{E},\ \hat{\mathbf{v}} \cdot \hat{\mathbf{n}}_{\hat{e}} \in \mathbb{P}_1(\hat{e})$$

where $\hat{e}$ represents a face of $\hat{E}$ and $\hat{\mathbf{n}}_{\hat{e}}$ the unit outward normal to $\hat{e}$. The multipoint flux approximation procedure requires on each face one velocity degree of freedom to be associated with each vertex thus requiring four degrees of freedom per face. Since $\mathbf{V}_h^*$ has only three degrees of freedom per face, it is augmented with one degree of freedom per face resulting in addition of six degrees of freedom per element. Since the properties of constant divergence, linear independence of the shape functions and continuity of the normal component across the element faces are to be preserved, six curl terms are added (see [50]) to $\mathbf{V}_h^*$. Let $\mathbf{V}_h \times W_h$ be the enhanced BDDF$_1$ spaces on hexahedra. On $\hat{E}$, these spaces are

$$\begin{aligned} \hat{\mathbf{V}}(\hat{E}) = \hat{\mathbf{V}}^*(\hat{E}) &+ r_2\, curl(0,0,\hat{x}^2\hat{z})^T + r_3\, curl(0,0,\hat{x}^2\hat{y}\hat{z})^T + s_2\, curl(\hat{x}\hat{y}^2,0,0)^T \\ &+ s_3\, curl(\hat{x}\hat{y}^2\hat{z}^2,0,0)^T + t_2\, curl(0,\hat{y}\hat{z}^2,0)^T + t_3\, curl(0,\hat{x}\hat{y}\hat{z}^2,0)^T \\ \hat{W}(\hat{E}) = \mathbb{P}_0(\hat{E}) & \end{aligned}$$

with the following properties

$$\hat{\nabla} \cdot \hat{\mathbf{V}}(\hat{E}) = \hat{W}(\hat{E}), \qquad \text{and} \qquad \forall \hat{\mathbf{v}} \in \hat{\mathbf{V}}(\hat{E}),\ \forall \hat{e} \subset \partial \hat{E},\ \hat{\mathbf{v}} \cdot \hat{\mathbf{n}}_{\hat{e}} \in \mathbb{Q}_1(\hat{e})$$

Since $dim\, \mathbb{Q}_1(\hat{e}) = 4$, the dimension of $\hat{\mathbf{V}}(\hat{E})$ is 24 as shown in Figure 2.2.

2.3.3 Weak form of Darcy law

We first rewrite the Darcy law (2.2) as

$$\frac{\mu}{\rho}\mathbf{K}^{-1}\mathbf{z} = -\nabla p + \rho\mathbf{g}$$

The weak form of the above is

$$\int\limits_{\Omega^f} \frac{\mu}{\rho}\mathbf{K}^{-1}\mathbf{z}\cdot\mathbf{v} = -\int\limits_{\Omega^f} \nabla p\cdot\mathbf{v} + \int\limits_{\Omega^f} \rho\mathbf{g}\cdot\mathbf{v} \tag{2.17}$$

We use the divergence theorem for the first term on RHS of (2.17) as follows

$$\begin{aligned}\int\limits_{\Omega^f} \nabla p\cdot\mathbf{v} &= \int\limits_{\Omega^f} \big(\nabla\cdot(p\mathbf{v}) - p\nabla\cdot\mathbf{v}\big)\\ &= \int\limits_{\partial\Omega^f} p\mathbf{v}\cdot\mathbf{n} - \int\limits_{\Omega^f} p\nabla\cdot\mathbf{v}\\ &= \int\limits_{\Gamma^f_D} g\mathbf{v}\cdot\mathbf{n} - \int\limits_{\Omega^f} p\nabla\cdot\mathbf{v} \end{aligned} \tag{2.18}$$

where we invoke $\mathbf{v}\cdot\mathbf{n} = 0$ on Γ^f_N. Substituting (2.18) in (2.17), we get

$$\int\limits_{\Omega^f} \frac{\mu}{\rho}\mathbf{K}^{-1}\mathbf{z}\cdot\mathbf{v} - \int\limits_{\Omega^f} p\nabla\cdot\mathbf{v} = \int\limits_{\Omega^f} \rho\mathbf{g}\cdot\mathbf{v} - \int\limits_{\Gamma^f_D} g\mathbf{v}\cdot\mathbf{n} \tag{2.19}$$

2.3.4 Discrete weak form of flow model

A subspace of $\mathbf{H}(div, \Omega^f)$ consisting of functions with vanishing normal trace on Γ^f_N is

$$\mathbf{V}(\Omega^f) \equiv \mathbf{H}(div, \Omega^f) \cap \left\{\mathbf{v} : \mathbf{v}\cdot\mathbf{n} = 0 \text{ on } \Gamma^f_N\right\}$$

The problem statement is : find $\mathbf{z}_h$, $\tilde{\mathbf{z}}_h \in \mathbf{V}_h$ and $p_h \in W_h$ such that

$$\begin{aligned}
&\sum_{E^f \in \mathscr{T}_h^f} \int_{E^f} \frac{\mu}{\rho} \mathbf{K}^{-1} \mathbf{z}_h \cdot \mathbf{v} - \sum_{E^f \in \mathscr{T}_h^f} \int_{E^f} p_h \nabla \cdot \mathbf{v} \\
&\qquad = - \sum_{E^f \in \mathscr{T}_h^f} \int_{\partial E^f \cap \Gamma_D^f} g \mathbf{v} \cdot \mathbf{n} + \sum_{E^f \in \mathscr{T}_h^f} \int_{E^f} \rho \mathbf{g} \cdot \mathbf{v} \qquad (2.20) \\
&\sum_{E^f \in \mathscr{T}_h^f} \int_{E^f} \phi^* \rho \, w + \sum_{E^f \in \mathscr{T}_h^f} \int_{E^f} \Delta t \nabla \cdot \mathbf{z}_h w \\
&\qquad = \sum_{E^f \in \mathscr{T}_h^f} \int_{E^f} \Delta t q w + \sum_{E^f \in \mathscr{T}_h^f} \int_{E^f} (\phi^* \rho)^n w \qquad (2.21)
\end{aligned}$$

where the finite dimensional subspaces $W_h \subset L^2(\Omega^f)$ and $\mathbf{V}_h \subset \mathbf{V}(\Omega^f)$ are

$$\begin{aligned}
W_h &\equiv \left\{ w : w|_{E^f} \in \mathbb{P}_0(E^f) \ \forall E^f \in \mathscr{T}_h^f \right\} \\
\mathbf{V}_h &\equiv \left\{ \mathbf{v} : \mathbf{v}|_{E^f} \leftrightarrow \hat{\mathbf{v}}|_{\hat{E}} : \hat{\mathbf{v}}|_{\hat{E}} \in \hat{\mathbf{V}}(\hat{E}) \ \forall E^f \in \mathscr{T}_h^f, \ \mathbf{v} \cdot \mathbf{n} = 0 \text{ on } \Gamma_N^f \right\}
\end{aligned}$$

$(\cdot)^n$ denotes the quantity $(\cdot)$ evaluated at the previous time level n while the remaining terms are evaluated at current time level $n+1$. Equations (2.20)-(2.21) are recast for the $(k+1)^{th}$ Newton iteration as

$$\begin{bmatrix} \mathbf{A} & -\mathbf{B} \\ \Delta t \, \mathbf{B}^T & \mathbf{C} \end{bmatrix} \left\{ \begin{array}{c} \delta^{(k)} \mathbf{Z}_h \\ \delta^{(k)} p_h \end{array} \right\} = \left\{ \begin{array}{c} R_1 \\ R_2 \end{array} \right\} \qquad (2.22)$$

where $\mathbf{Z}_h$ are the flux degrees of freedom and p_h are the pressure degrees of freedom respectively and the submatrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are

$$\begin{aligned}
\mathbf{A} &= \sum_{E^f \in \mathscr{T}_h^f} \int_{E^f} \frac{\mu}{\rho^k} \mathbf{K}^{-1} \mathbf{v}_j \cdot \mathbf{v}_i \\
\mathbf{B} &= \sum_{E^f \in \mathscr{T}_h^f} \int_{E^f} \nabla \cdot \mathbf{v}_i \, w_j
\end{aligned}$$

$$\mathbf{C} = \sum_{E^f \in \mathscr{T}_h^f} \int_{E^f} \rho^k {\phi^*}^k c + \sum_{E^f \in \mathscr{T}_h^f} \int_{E^f} \rho^k \frac{\partial \phi^*}{\partial p}|_k \tag{2.23}$$

and the residuals R_1 and R_2 are

$$\begin{aligned} R_1 = &- \sum_{E^f \in \mathscr{T}_h^f} \int_{E^f} \mathbf{K}^{-1} \tilde{\mathbf{z}}_h^k \cdot \mathbf{v} + \sum_{E^f \in \mathscr{T}_h^f} \int_{E^f} p_h^k \nabla \cdot \mathbf{v} \\ &- \sum_{E^f \in \mathscr{T}_h^f} \int_{\partial E^f \cap \Gamma_D^f} g \mathbf{v} \cdot \mathbf{n} + \sum_{E^f \in \mathscr{T}_h^f} \int_{E^f} \rho^k \mathbf{g} \cdot \mathbf{v} \end{aligned} \tag{2.24}$$

$$\begin{aligned} R_2 = &- \sum_{E^f \in \mathscr{T}_h^f} \int_{E^f} {\phi^*}^k \rho^k w - \sum_{E^f \in \mathscr{T}_h^f} \int_{E^f} \Delta t \nabla \cdot \mathbf{z}_h^k w \\ &+ \sum_{E^f \in \mathscr{T}_h^f} \int_{E^f} \Delta t q w + \sum_{E^f \in \mathscr{T}_h^f} \int_{E^f} (\phi^* \rho)^n w \end{aligned} \tag{2.25}$$

where $(\cdot)^k$ is the quantity $(\cdot)$ evaluated at the k^{th} Newton iteration at the current time level $n+1$ and $\delta^{(k)}$ is the change in $(\cdot)$ over the $(k+1)^{th}$ iteration. The submatrices $\mathbf{A}$ and $\mathbf{B}$ have off-diagonal components while the submatrix $\mathbf{C}$ is diagonal. The evaluation of $\frac{\partial \phi^*}{\partial p}|_k$ in (2.23) along with the evolution law for ϕ^* during the flow solve are presented in the module 2.3.7. A quadrature rule (Ingram et al. [50]) is used to evaluate the submatrix $\mathbf{A}$. Elimination of the flux degrees of freedom from (2.22) results in a 27 point stencil where the pressure in each element E^f is coupled with pressures in all 27 elements that share a vertex with E^f. The resulting system is solved for $\delta^{(k)} p_h$ and the k^{th} iterate is obtained as

$$p_h^{k+1} = p_h^k + \delta^{(k)} p_h \tag{2.26}$$

2.3.5 Discrete weak form of linear momentum balance

An infinite dimensional space $\mathbf{U}(\Omega^p)$ in Ω^p is defined as

$$\mathbf{U}(\Omega^p) \equiv \left\{ \mathbf{q} = (u, v, w) : u, v, w \in H^1(\Omega^p), \mathbf{q} = \mathbf{0} \text{ on } \Gamma^p_D \right\}$$

The weak form of the linear momentum balance (2.6) is given by

$$\int\limits_{\Omega^p} (\nabla \cdot \boldsymbol{\sigma}) \cdot \mathbf{q} + \int\limits_{\Omega^p} \mathbf{f} \cdot \mathbf{q} = 0 \qquad (\forall\ \mathbf{q} \in \mathbf{U}(\Omega^p)) \tag{2.27}$$

where $\mathbf{U}(\Omega^p)$ is given by

$$\mathbf{U}(\Omega^p) \equiv \left\{ \mathbf{q} = (u, v, w) : u, v, w \in H^1(\Omega^p), \mathbf{q} = \mathbf{0} \text{ on } \Gamma^p_D \right\}$$

We know from tensor calculus that

$$(\nabla \cdot \boldsymbol{\sigma}) \cdot \mathbf{q} \equiv \nabla \cdot (\boldsymbol{\sigma}\mathbf{q}) - \boldsymbol{\sigma} : \nabla \mathbf{q} \tag{2.28}$$

Further, using the divergence theorem and the symmetry of $\boldsymbol{\sigma}$, we arrive at

$$\int\limits_{\Omega^p} \nabla \cdot (\boldsymbol{\sigma}\mathbf{q}) \equiv \int\limits_{\partial\Omega^p} \mathbf{q} \cdot (\boldsymbol{\sigma}\mathbf{n}) \tag{2.29}$$

We decompose $\nabla \mathbf{q}$ into a symmetric part $(\nabla \mathbf{q})_s \equiv \frac{1}{2}\big(\nabla \mathbf{q} + (\nabla \mathbf{q})^T\big) \equiv \boldsymbol{\epsilon}(\mathbf{q})$ and skew-symmetric part $(\nabla \mathbf{q})_{ss}$ and note that the contraction between a symmetric and skew-symmetric tensor is zero to obtain

$$\boldsymbol{\sigma} : \nabla \mathbf{q} \equiv \boldsymbol{\sigma} : (\nabla \mathbf{q})_s + \cancelto{0}{\boldsymbol{\sigma} : (\nabla \mathbf{q})_{ss}} = \boldsymbol{\sigma} : \boldsymbol{\epsilon}(\mathbf{q}) \tag{2.30}$$

From (5.7), (5.8), (5.9) and (5.10), we get

$$\int\limits_{\partial\Omega^p} (\boldsymbol{\sigma}\mathbf{n}) \cdot \mathbf{q} - \int\limits_{\Omega^p} \boldsymbol{\sigma} : \boldsymbol{\epsilon}(\mathbf{q}) + \int\limits_{\Omega^p} \mathbf{f} \cdot \mathbf{q} = 0$$

which, after invoking the boundary conditions $\boldsymbol{\sigma}\mathbf{n} = \mathbf{t}$ on Γ_N^p and $\mathbf{q} = 0$ on Γ_D^p results in

$$\int\limits_{\Gamma_N^p} \mathbf{t}\cdot\mathbf{q} - \int\limits_{\Omega^p} \boldsymbol{\sigma} : \boldsymbol{\epsilon}(\mathbf{q}) + \int\limits_{\Omega^p} \mathbf{f}\cdot\mathbf{q} = 0 \tag{2.31}$$

Substituting $\boldsymbol{\sigma} = \mathbb{D}\boldsymbol{\epsilon} - \alpha p\mathbf{I}$ in (4.21), we get

$$\int\limits_{\Gamma_N^p} \mathbf{t}\cdot\mathbf{q} - \int\limits_{\Omega^p} (\mathbb{D}\boldsymbol{\epsilon} - \alpha p\mathbf{I}) : \boldsymbol{\epsilon}(\mathbf{q}) + \int\limits_{\Omega^p} \mathbf{f}\cdot\mathbf{q} = 0$$

and the discrete weak form after substituting $\mathbf{f} = \rho\phi\mathbf{g} + \rho_r(1-\phi)\mathbf{g}$ is finally obtained as: find $\mathbf{u}_H \in \mathbf{U}_H$ such that

$$\begin{aligned}\sum_{E^f\in\mathscr{T}_H^p}\int\limits_{E^p} \boldsymbol{\epsilon}(\mathbf{q}) : \mathbb{D}\boldsymbol{\epsilon}(\mathbf{u}_H) = & \sum_{E^f\in\mathscr{T}_H^p}\int\limits_{\Omega^p} \boldsymbol{\epsilon}(\mathbf{q}) : \alpha p\mathbf{I} \\ & + \sum_{E^f\in\mathscr{T}_H^p}\int\limits_{E^p} \mathbf{q}\cdot(\rho\phi\mathbf{g} + \rho_r(1-\phi)\mathbf{g}) + \sum_{E^f\in\mathscr{T}_H^p}\int\limits_{\partial E^p\cap\Gamma_N^p} \mathbf{q}\cdot\mathbf{t}\end{aligned} \tag{2.32}$$

where the finite dimensional subspace $\mathbf{U}_H \subset \mathbf{U}(\Omega^p)$ is given by

$$\mathbf{U}_H = \left\{\mathbf{q} = (u, v, w) : u|_{E^p}, v|_{E^p}, w|_{E^p} \in \mathbb{Q}_1(E^p)\ \forall E^p \in \mathscr{T}_H^p, \mathbf{q} = \mathbf{0} \text{ on } \Gamma_D^p\right\}$$

The discrete displacements $\mathbf{u}_H$ and the corresponding strain tensor $\boldsymbol{\epsilon}(\mathbf{u}_H)$ are written in terms of the nodal displacement degrees of freedom represented by $\mathcal{U}$ as

$$\mathbf{u}_H = \sum_{E^p\in\mathscr{T}_H^p} \mathbf{N}\mathcal{U} \tag{2.33}$$

$$\boldsymbol{\epsilon}(\mathbf{u}_H) = \sum_{E^p\in\mathscr{T}_H^p} \mathbf{B}\mathcal{U} \tag{2.34}$$

where **N** is the shape function matrix and **B** is the strain-displacement interpolation matrix. Equations (2.32), (2.33) and (2.34) eventually lead to the following system of equations

$$\mathbf{K}\mathcal{U} = \mathbf{F} \tag{2.35}$$

where **K** and **F** are refered to as the global stiffness matrix and global force vector respectively and are obtained as

$$\mathbf{K} = \sum_{E^p \in \mathscr{T}_H^p} \int_{E^p} \mathbf{B}^T \mathbb{D} \mathbf{B} \tag{2.36}$$

$$\mathbf{F} = \sum_{E^p \in \mathscr{T}_H^p} \int_{E^p} \mathbf{B}^T \alpha p \mathbf{I} + \sum_{E^p \in \mathscr{T}_H^p} \int_{E^p} \mathbf{N}^T (\rho \phi \mathbf{g} + \rho_r (1 - \phi) \mathbf{g}) + \sum_{E^p \in \mathscr{T}_H^p} \int_{\partial E^p \cap \Gamma_N^p} \mathbf{N}^T \mathbf{t} \tag{2.37}$$

To simplify the computations, (2.35) is recast in compact engineering notation (see Hughes [48]) wherein stresses $\boldsymbol{\sigma}$, strains $\boldsymbol{\epsilon}$ and identity tensor **I** are represented as vectors and fourth order tensor $\mathbb{D}$ is represented as a second order tensor. The matrices **N** and **B** are also recast appropriately.

2.3.6 Evaluation of integrals for the poromechanical solve using quadrature

The trilinear function $F_E^p(\hat{\mathbf{r}}) : \hat{\mathbf{r}} \mapsto \mathbf{r}^p$ which we have already defined in (2.14) can also be written as

$$F_E^p(\hat{\mathbf{r}}) : \hat{\mathbf{r}} \mapsto \mathbf{r}^p = \sum_{i=1}^{8} \hat{N}_i(\hat{\mathbf{r}}) \mathbf{r}_i^p$$

where $\hat{N}_i(\hat{\mathbf{r}})$ are refered to as the shape functions and satisfy the property that $\hat{N}_i(\hat{\mathbf{r}}_j) = \delta_{ij}$, $i, j = 1, ..., 8$ where δ_{ij} is the Kronecker delta, and are given as

$$\begin{aligned}
\hat{N}_1 &= (1-\hat{x})(1-\hat{y})(1-\hat{z}) \\
\hat{N}_2 &= \hat{x}(1-\hat{y})(1-\hat{z}) \\
\hat{N}_3 &= \hat{x}\hat{y}(1-\hat{z}) \\
\hat{N}_4 &= (1-\hat{x})\hat{y}(1-\hat{z}) \\
\hat{N}_5 &= (1-\hat{x})(1-\hat{y})\hat{z} \\
\hat{N}_6 &= \hat{x}(1-\hat{y})\hat{z} \\
\hat{N}_7 &= \hat{x}\hat{y}\hat{z} \\
\hat{N}_8 &= (1-\hat{x})\hat{y}\hat{z}
\end{aligned}$$

The shape function matrix is given as

$$\begin{aligned}
\hat{\mathbf{N}} &= \begin{bmatrix} \bar{\mathbf{N}}_1 & \bar{\mathbf{N}}_2 & \bar{\mathbf{N}}_3 & \bar{\mathbf{N}}_4 & \bar{\mathbf{N}}_5 & \bar{\mathbf{N}}_6 & \bar{\mathbf{N}}_7 & \bar{\mathbf{N}}_8 \end{bmatrix}_{3\times 24} \\
\bar{\mathbf{N}}_i &= \begin{bmatrix} \hat{N}_i & 0 & 0 \\ 0 & \hat{N}_i & 0 \\ 0 & 0 & \hat{N}_i \end{bmatrix}_{3\times 3} \qquad (\forall\ i = 1, .., 8)
\end{aligned}$$

The jacobian of $F_E^p(\hat{\mathbf{r}})$ is the given as

$$DF_E^p(\hat{\mathbf{r}}) \equiv \begin{bmatrix} \frac{\partial x}{\partial \hat{x}} & \frac{\partial x}{\partial \hat{y}} & \frac{\partial x}{\partial \hat{z}} \\ \frac{\partial y}{\partial \hat{x}} & \frac{\partial y}{\partial \hat{y}} & \frac{\partial y}{\partial \hat{z}} \\ \frac{\partial z}{\partial \hat{x}} & \frac{\partial z}{\partial \hat{y}} & \frac{\partial z}{\partial \hat{z}} \end{bmatrix} = \begin{bmatrix} \sum\limits_{i=1}^{8} \frac{\partial \hat{N}_i}{\partial \hat{x}} x_i^p & \sum\limits_{i=1}^{8} \frac{\partial \hat{N}_i}{\partial \hat{y}} x_i^p & \sum\limits_{i=1}^{8} \frac{\partial \hat{N}_i}{\partial \hat{z}} x_i^p \\ \sum\limits_{i=1}^{8} \frac{\partial \hat{N}_i}{\partial \hat{x}} y_i^p & \sum\limits_{i=1}^{8} \frac{\partial \hat{N}_i}{\partial \hat{y}} y_i^p & \sum\limits_{i=1}^{8} \frac{\partial \hat{N}_i}{\partial \hat{z}} y_i^p \\ \sum\limits_{i=1}^{8} \frac{\partial \hat{N}_i}{\partial \hat{x}} z_i^p & \sum\limits_{i=1}^{8} \frac{\partial \hat{N}_i}{\partial \hat{y}} z_i^p & \sum\limits_{i=1}^{8} \frac{\partial \hat{N}_i}{\partial \hat{z}} z_i^p \end{bmatrix}$$

The strain-displacement interpolation matrix is given as

$$\hat{\mathbf{B}} = \begin{bmatrix} \hat{\mathbf{B}}_1 & \hat{\mathbf{B}}_2 & \hat{\mathbf{B}}_3 & \hat{\mathbf{B}}_4 & \hat{\mathbf{B}}_5 & \hat{\mathbf{B}}_6 & \hat{\mathbf{B}}_7 & \hat{\mathbf{B}}_8 \end{bmatrix}_{6\times 24}$$

$$\hat{\mathbf{B}}_i = \begin{bmatrix} \frac{\partial \hat{\mathbf{N}}_i}{\partial x} & 0 & 0 \\ 0 & \frac{\partial \hat{\mathbf{N}}_i}{\partial y} & 0 \\ 0 & 0 & \frac{\partial \hat{\mathbf{N}}_i}{\partial z} \\ \frac{\partial \hat{\mathbf{N}}_i}{\partial y} & \frac{\partial \hat{\mathbf{N}}_i}{\partial x} & 0 \\ \frac{\partial \hat{\mathbf{N}}_i}{\partial z} & 0 & \frac{\partial \hat{\mathbf{N}}_i}{\partial x} \\ 0 & \frac{\partial \hat{\mathbf{N}}_i}{\partial z} & \frac{\partial \hat{\mathbf{N}}_i}{\partial y} \end{bmatrix}_{6\times 3} \qquad (\forall\ i = 1,..,8)$$

and the derivatives of the shape functions are obtained using

$$\begin{bmatrix} \frac{\partial \hat{\mathbf{N}}_i}{\partial x} \\ \frac{\partial \hat{\mathbf{N}}_i}{\partial y} \\ \frac{\partial \hat{\mathbf{N}}_i}{\partial z} \end{bmatrix} = (DF_E^p)^{-T} \begin{bmatrix} \frac{\partial \hat{\mathbf{N}}_i}{\partial \hat{x}} \\ \frac{\partial \hat{\mathbf{N}}_i}{\partial \hat{y}} \\ \frac{\partial \hat{\mathbf{N}}_i}{\partial \hat{z}} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{8} \frac{\partial \hat{N}_i}{\partial \hat{x}} x_i^p & \sum_{i=1}^{8} \frac{\partial \hat{N}_i}{\partial \hat{y}} x_i^p & \sum_{i=1}^{8} \frac{\partial \hat{N}_i}{\partial \hat{z}} x_i^p \\ \sum_{i=1}^{8} \frac{\partial \hat{N}_i}{\partial \hat{x}} y_i^p & \sum_{i=1}^{8} \frac{\partial \hat{N}_i}{\partial \hat{y}} y_i^p & \sum_{i=1}^{8} \frac{\partial \hat{N}_i}{\partial \hat{z}} y_i^p \\ \sum_{i=1}^{8} \frac{\partial \hat{N}_i}{\partial \hat{x}} z_i^p & \sum_{i=1}^{8} \frac{\partial \hat{N}_i}{\partial \hat{y}} z_i^p & \sum_{i=1}^{8} \frac{\partial \hat{N}_i}{\partial \hat{z}} z_i^p \end{bmatrix}^{-T} \begin{bmatrix} \frac{\partial \hat{\mathbf{N}}_i}{\partial \hat{x}} \\ \frac{\partial \hat{\mathbf{N}}_i}{\partial \hat{y}} \\ \frac{\partial \hat{\mathbf{N}}_i}{\partial \hat{z}} \end{bmatrix}$$

The elasticity tensor for the isotropic porous solid is

$$\mathbb{D} = \begin{bmatrix} \lambda + 2G & \lambda & \lambda & 0 & 0 & 0 \\ \lambda & \lambda + 2G & \lambda & 0 & 0 & 0 \\ \lambda & \lambda & \lambda + 2G & 0 & 0 & 0 \\ 0 & 0 & 0 & G & 0 & 0 \\ 0 & 0 & 0 & 0 & G & 0 \\ 0 & 0 & 0 & 0 & 0 & G \end{bmatrix}_{6\times 6}$$

where λ is the Lame parameter corresponding to the drained response of the porous specimen and G is the shear modulus of porous specimen. Denoting the determinant of $DF_E^p(\hat{\mathbf{r}})$ by J_E^p, the integrals leading upto the element stiffness matrix and force vectors are given as

$$\left.\begin{aligned}
\int\limits_{E^p} \mathbf{N}^T\mathbf{f} &\equiv \int\limits_{\hat{E}} \hat{\mathbf{N}}^T\mathbf{f}J_E^p \equiv \sum_{g\in\mathscr{G}}(\hat{\mathbf{N}}^T\mathbf{f}J_E^p)|_g w_g \\
\int\limits_{E^p} \mathbf{B}^T\alpha p\mathbf{I} &\equiv \int\limits_{\hat{E}} \hat{\mathbf{B}}^T\alpha p\mathbf{I}J_E^p \equiv \sum_{g\in\mathscr{G}}(\hat{\mathbf{B}}^T\alpha p\mathbf{I}J_E^p)|_g w_g \\
\int\limits_{E^p} \mathbf{B}^T\mathbb{D}\mathbf{B} &\equiv \int\limits_{\hat{E}} \hat{\mathbf{B}}^T\mathbb{D}\hat{\mathbf{B}}J_E^p \equiv \sum_{g\in\mathscr{G}}(\hat{\mathbf{B}}^T\mathbb{D}\hat{\mathbf{B}}J_E^p)|_g w_g \\
\int\limits_{\partial E^p\cap\Gamma_N^p} \mathbf{N}^T\mathbf{t} &\equiv \int\limits_{\partial\hat{E}\cap\Gamma_N^p} \hat{\mathbf{N}}^T\mathbf{t}\,J_E^p \equiv \sum_{g\in\mathscr{G}\cap\partial\hat{E}}(\hat{\mathbf{N}}^T\mathbf{t}\,J_E^p)|_g w_g
\end{aligned}\right\}\tag{2.38}$$

where $\mathscr{G}$ is the set (of cardinality 8) of quadrature points, $(\cdot)|_g$ is the quantity $(\cdot)$ evaluated at the quadrature point $g\in\mathscr{G}$, $\mathbf{I} = \begin{bmatrix}1 & 1 & 1 & 0 & 0 & 0\end{bmatrix}^T$ and w_g is the weight associated with a quadrature point $g\in\mathscr{G}$. In lieu of (2.35)-(2.37) and (2.38), we write

$$\mathbf{K}\mathcal{U} = \mathbf{F} \tag{2.39}$$

where the global stiffness matrix $\mathbf{K}$ and global force vector $\mathbf{F}$

$$\mathbf{K} = \sum_{E^p\in\mathscr{T}_H^p}\sum_{g\in\mathscr{G}}(\hat{\mathbf{B}}^T\mathbb{D}\hat{\mathbf{B}}J_E^p)|_g w_g \tag{2.40}$$

$$\begin{aligned}
\mathbf{F} = &\sum_{E^p\in\mathscr{T}_H^p}\sum_{g\in\mathscr{G}}(\hat{\mathbf{B}}^T\alpha p\mathbf{I}J_E^p)|_g w_g + \sum_{E^p\in\mathscr{T}_H^p}\sum_{g\in\mathscr{G}}(\hat{\mathbf{N}}^T\mathbf{f}J_E^p)|_g w_g \\
&+ \sum_{E^p\in\mathscr{T}_H^p}\sum_{g\in\mathscr{G}\cap\partial\hat{E}}(\hat{\mathbf{N}}^T\mathbf{t}\,J_E^p)|_g w_g
\end{aligned}\tag{2.41}$$

are obtained by assembling the element stiffness matrix and element force vectors. The reader is refered to Zienkiewicz et al. [101] for the details of the assembly process. The set $\mathscr{G}$ of quadrature points with associated weights is

$$\begin{aligned}
g_1 &\equiv \left(\frac{1}{2}\left(1-\frac{\sqrt{3}}{3}\right), \frac{1}{2}\left(1-\frac{\sqrt{3}}{3}\right), \frac{1}{2}\left(1-\frac{\sqrt{3}}{3}\right)\right)\\
g_2 &\equiv \left(\frac{1}{2}\left(1+\frac{\sqrt{3}}{3}\right), \frac{1}{2}\left(1-\frac{\sqrt{3}}{3}\right), \frac{1}{2}\left(1-\frac{\sqrt{3}}{3}\right)\right)\\
g_3 &\equiv \left(\frac{1}{2}\left(1+\frac{\sqrt{3}}{3}\right), \frac{1}{2}\left(1+\frac{\sqrt{3}}{3}\right), \frac{1}{2}\left(1-\frac{\sqrt{3}}{3}\right)\right)\\
g_4 &\equiv \left(\frac{1}{2}\left(1-\frac{\sqrt{3}}{3}\right), \frac{1}{2}\left(1+\frac{\sqrt{3}}{3}\right), \frac{1}{2}\left(1-\frac{\sqrt{3}}{3}\right)\right)\\
g_5 &\equiv \left(\frac{1}{2}\left(1-\frac{\sqrt{3}}{3}\right), \frac{1}{2}\left(1-\frac{\sqrt{3}}{3}\right), \frac{1}{2}\left(1+\frac{\sqrt{3}}{3}\right)\right)\\
g_6 &\equiv \left(\frac{1}{2}\left(1+\frac{\sqrt{3}}{3}\right), \frac{1}{2}\left(1-\frac{\sqrt{3}}{3}\right), \frac{1}{2}\left(1+\frac{\sqrt{3}}{3}\right)\right)\\
g_7 &\equiv \left(\frac{1}{2}\left(1+\frac{\sqrt{3}}{3}\right), \frac{1}{2}\left(1+\frac{\sqrt{3}}{3}\right), \frac{1}{2}\left(1+\frac{\sqrt{3}}{3}\right)\right)\\
g_8 &\equiv \left(\frac{1}{2}\left(1-\frac{\sqrt{3}}{3}\right), \frac{1}{2}\left(1+\frac{\sqrt{3}}{3}\right), \frac{1}{2}\left(1+\frac{\sqrt{3}}{3}\right)\right)\\
w_g &= \frac{1}{8} \qquad \forall\ g \in \mathscr{G}
\end{aligned}$$

2.3.7 Lagrangian porosity update during the flow solve with the fixed mean stress constraint

The Eulerian porosity variation in a deformable porous medium is approximated as (Geertsma [36], Brown and Korringa [17])

$$\delta\phi = \left(\frac{\alpha-\phi}{K_b}\right)(\delta\sigma_v + \delta p) \tag{2.42}$$

Imposing the fixed mean stress constraint $\delta\sigma_v = 0$ in (2.42) results in

$$\delta\phi = \left(\frac{\alpha-\phi}{K_b}\right)\delta p \tag{2.43}$$

Using the relation for the mean stress $\sigma_v = K_b\bar{\epsilon} - \alpha p$ and the fixed mean stress constraint $\delta\sigma_v = 0$ results in

$$\delta\bar{\epsilon} = \frac{\alpha}{K_b}\delta p$$

which implies that the evolution law for volumetric strain is given by

$$\epsilon^{k+1} = \epsilon^{k} + \frac{\alpha}{K_b}\delta^{(k)}p$$

In lieu of (2.43) and the relationship $\phi^* = \phi(1+\bar{\epsilon})$, the Lagrangian porosity variation is given by

$$\delta\phi^* = \phi\delta\bar{\epsilon} + (1+\bar{\epsilon})\delta\phi = \left((\phi + 1 + \bar{\epsilon})\frac{\alpha}{K_b} - \frac{\phi^*}{K}\right)\delta p$$

which implies that the evolution law for Lagrangian porosity is given by

$$\phi^{*^{k+1}} = \phi^{*k} + \overbrace{\left((\phi^k + 1 + \bar{\epsilon}^k)\frac{\alpha}{K_b} - \frac{\phi^{*^k}}{K_b}\right)}^{\frac{\partial\phi^*}{\partial p}|_k}\delta^{(k)}p \tag{2.44}$$

2.3.8 Lagrangian porosity update during the poromechanical solve

Invoking (2.42) again and using the relation $\delta\sigma_v = K_b\delta\bar{\epsilon} - \alpha\delta p$, we get

$$\delta\phi = (\alpha - \phi)\delta\bar{\epsilon} + \frac{(\alpha - \phi)(1-\alpha)}{K_b}\delta p \tag{2.45}$$

In lieu of (2.45) and the relationship $\phi^* = \phi(1+\bar{\epsilon})$, the Lagrangian porosity variation is given by

$$\delta\phi^* = \phi\delta\bar{\epsilon} + (1+\bar{\epsilon})\delta\phi = \phi\delta\bar{\epsilon} + (1+\bar{\epsilon})\left((\alpha-\phi)\delta\bar{\epsilon} + \frac{(\alpha - \phi)(1-\alpha)}{K_b}\delta p\right)$$

$$= (\alpha + \bar{\epsilon}(\alpha - \phi))\delta\bar{\epsilon} + (1 + \bar{\epsilon})\frac{(\alpha - \phi)(1 - \alpha)}{K_b}\delta p \tag{2.46}$$

Following the arguments of Coussy [24], linear poroelasticity consists in setting the tangents $(\alpha + \bar{\epsilon}(\alpha - \phi))$ and $(1 + \bar{\epsilon})\frac{(\alpha-\phi)(1-\alpha)}{K_b}$ in (2.46) as constants. Hence (2.46) is rewritten as

$$\delta\phi^* = (\alpha + \bar{\epsilon}_0(\alpha - \phi_0))\delta\bar{\epsilon} + (1 + \bar{\epsilon}_0)\frac{(\alpha - \phi_0)(1 - \alpha)}{K_b}\delta p$$

which can be integrated from initial time to obtain

$$\phi^{*^m} = \phi_0 + (\alpha + \bar{\epsilon}_0(\alpha - \phi_0))(\bar{\epsilon}^m - \bar{\epsilon}_0) + (1 + \bar{\epsilon}_0)\frac{(\alpha - \phi_0)(1 - \alpha)}{K_b}(p^m - p_0)$$

where $(\cdot)^m$ refers to $(\cdot)$ evaluated at the m^{th} coupling iteration. In lieu of the small strain assumption (2.9), the $O(\epsilon^2)$ terms are neglected to obtain

$$\phi^{*^m} = \phi_0 + \alpha(\bar{\epsilon}^m - \bar{\epsilon}_0) + (1 + \bar{\epsilon}_0)\frac{(\alpha - \phi_0)(1 - \alpha)}{K_b}(p^m - p_0)$$

which, under the initial condition (2.11) lending to $\bar{\epsilon}_0 = 0$, is finally written as

$$\phi^{*^m} = \phi_0 + \alpha\bar{\epsilon}^m + \frac{(\alpha - \phi_0)(1 - \alpha)}{K_b}(p^m - p_0) \tag{2.47}$$

2.4 The two-grid staggered solution algorithm

The flowchart of the algorithm is provided in Figure 2.3. The algorithm solves the flow system on a fine mesh while freezing the mean stress of the porous solid followed by the poromechanics system on a coarse mesh. The pressures at the end of the flow solve along with updated fluid densities and Lagrangian porosities are *upscaled* and these upscaled quantities contribute to the force vector in the poromechanics system. The poromechanics system is then solved for nodal displacements which are post-processed to obtain volumetric strains. These volumetric strains along with upscaled pressures are then used to update the Lagrangian porosities in the poromechanical grid. These coarse grid Lagrangian porosities are then *downscaled* onto the flow grid. This two-step procedure of nonlinear flow solve followed by linear poromechanics solve is continued until a convergence criterion (evaluated at the coarse scale poromechanics grid) is met. Once this criterion is met, we march forward to the next time step as shown in Algorithm 1. The design of the upscaling and downscaling operators is the subject of Chapter 3

2.4.1 Convergence criterion for coupling iterations

The stopping criterion for coupling iterations is designed based on the requirement that the change in mean stress across coupling iterations in each time step reduces monotonically from one coupling iteration to the next. This is also the notion of the contraction map that we invoke for the convergence analysis of the algorithm in Chapter 4. We already know that one coupling iteration consists of one flow solve and one poromechanics solve. We also know that the fixed stress split strategy imposes a fixed mean stress during the flow solve in every coupling iteration. This implies that the mean stress only changes during the poromechanics solve in every coupling iteration. This automatically implies that the change in mean stress across one coupling iteration is identical to the change in mean stress during the poromechanics solve in that coupling iteration. As a result, the stopping criterion is to be designed based on the requirement that the change in mean stress across poromechanics solves reduces monotonically from one coupling iteration to the next. Now, let us look at the expression (2.42) for porosity variation in deformable porous medium given by

$$\delta\phi = \left(\frac{\alpha - \phi}{K_b}\right)(\delta\sigma_v + \delta p)$$

and let us write it in two forms

$$\begin{aligned}(\delta\phi)|_f &= \left(\frac{\alpha - \phi}{K_b}\right)((\delta\sigma_v)|_f + (\delta p)|_f)\\(\delta\phi)|_c &= \left(\frac{\alpha - \phi}{K_b}\right)((\delta\sigma_v)|_c + (\delta p)|_c)\end{aligned}$$

where $\delta(\cdot)|_f$ is the change in quantity $(\cdot)$ during the flow solve in a coupling iteration and $\delta(\cdot)|_c$ is the change in quantity $(\cdot)$ over the entire coupling iteration. Now the fixed mean stress constraint implies that $(\delta\sigma_v)|_f = 0$ and since the pressure does not change during the poromechanics solve, $(\delta p)|_f = (\delta p)|_c$. As a result, we have

$$(\delta\phi)|_c - (\delta\phi)|_f = \left(\frac{\alpha - \phi}{K_b}\right)(\delta\sigma_v)|_c \equiv \left(\frac{\alpha - \phi}{K_b}\right)(\delta\sigma_v)|_m$$

where $(\delta\sigma_v)|_m$ is the change in mean stress during the poromechanics solve in the coupling iteration. In lieu of the above, the stopping criterion is designed based on the change in porosity during the poromechanics solve in every coupling iteration. Since we calculate the Lagrangian porosity updates after every flow and poromechanics solve, the stopping criterion is based in the requirement that the change in Lagrangian porosity during the poromechanics solve in every coupling iteration reduces monotonically from one coupling iteration to the next. Now the change in Lagrangian porosity during the poromechanics solve in the m^{th} coupling iteration in every time step is given by

$$\phi^{*^{m}} - \phi^{*^{k+1,m}}$$

where $\phi^{*^{m}}$ is the Lagrangian porosity at the end of the m^{th} coupling iteration, $\phi^{*^{k+1,m}}$ is the Lagrangian porosity at the end of the flow solve in the m^{th} coupling iteration and $k+1$ is the number of Newton iterations required to solve the nonlinear flow problem. This quantity is normalized by the current value of Lagrangian porosity i.e. $\phi^{*^{m}}$ as follows

$$\frac{\phi^{*^{m}} - \phi^{*^{k+1,m}}}{\phi^{*^{m}}}$$

and the stopping criterion is based on the requirement that the maximum norm of this quantity is less than a certain pre-specified tolerance as follows

$$\left\| \frac{\phi^{*^{m}} - \phi^{*^{k+1,m}}}{\phi^{*^{m}}} \right\|_{\infty} \leq \mathrm{TOL}$$

We shall revisit this stopping criterion but from the standpoint of fluid content in chapter 5, for the case of linearized single phase flow coupled with anisotropic poroelasticity. The concept of fluid content itself shall be introduced in chapter 4.

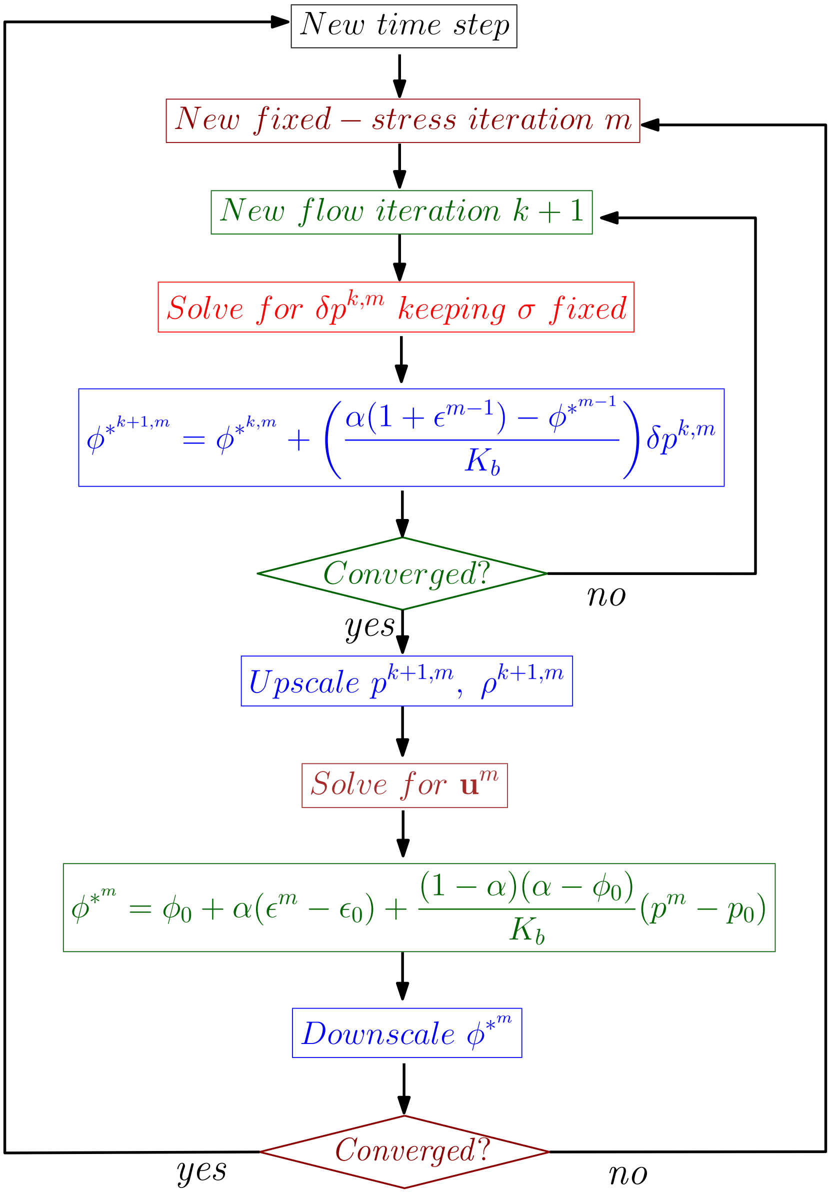

Figure 2.3: Flowchart for two grid fixed-stress split iterative scheme for single phase flow coupled with linear poromechanics and $(\cdot)^{k,m}$ denotes the quantity $(\cdot)$ evaluated at the k^{th} flow iteration within the m^{th} coupling iteration and $(\cdot)^m$ denotes the quantity $(\cdot)$ evaluated at the m^{th} coupling iteration.

Algorithm 1 two grid fixed-stress split iterative scheme for slightly compressible single phase flow coupled with linear poromechanics

```
For each time $t^n$
repeat                         ▷ Fixed-stress split iterative coupling loop (index $m$)
    repeat                                              ▷ Flow loop (index $k+1$)
        Solve for $\delta p_h^{k,m}, \delta \mathbf{z}_h^{k,m}$ using 2.22
        Update $p_h^{k+1,m}$, $\rho^{k+1,m}$ and $\phi^{*^{k+1,m}}$ using 2.26, 2.3 and 2.44 respectively
    until Stopping criterion, use 2.25
            $\|R_2\|_\infty \leq \mathrm{TOL}_1 \ \wedge \ \|R_2\|_1 \leq \mathrm{TOL}_2 \quad \mathrm{TOL}_1, \mathrm{TOL}_2 > 0$
    Upscale $p_h^{k+1,m}$, $\rho^{k+1,m}$ and $\phi^{*^{k+1,m}}$
    Solve for $\mathbf{u}_H^m$ using 2.35                          ▷ Mechanics solve
    Update $\phi^{*^m}$ using 2.47
    Downscale $\phi^{*^m}$
until Stopping criterion
            $\left\| \frac{\phi^{*^m} - \phi^{*^{k+1,m}}}{\phi^{*^m}} \right\|_\infty \leq \mathrm{TOL}_3, \quad \mathrm{TOL}_3 > 0$
Increment $t^n \rightarrow t^{n+1}$
```

2.5 Summary

We present the model equations for slightly compressible single phase flow coupled with linear poromechanics for flow through porous media. We then present the discrete variational statements for the two-grid staggered solution algorithm. The two-grid algorithm connotates that the model equations for flow and poromechanics are solved on different finite element grids. Flow is solved on a fine mesh using a multipoint flux mixed finite element method and poromechanics is solved on a coarse mesh using a conforming Galerkin finite element method. The notion of staggering lies in solving the flow and poromechanics equations sequentially in a coupling iteration instead of solving the coupled system of equations monolithically. The system solve at a particular time step can consist of multiple coupling iterations depending on when the stopping criterion is achieved. In essence, the staggering imposes a constraint on the system during the flow solve thereby working as a predictor to the state of the system followed by the poromechanics solve which works as a corrector to the state of the system at the end of the coupling iteration.

Chapter 3

Construction of the up and downscaling operators and verification of the two-grid algorithm

3.1 Parallelism involved in IPARS

We first comment on the MPI based parallelism involved in IPARS. The elements of the geomechanical domain are first divided areally $(y - z)$ among MPI processes. Since the flow domain is a subset of the geomechanical domain, the areal extent of the flow domain would be smaller than that of the geomechanical domain. In lieu of that, not all processes would occupy elements of the flow domain. We need to identify the processes that would occupy flow elements, and then carve out a MPI subcommunicator with only those processes involved. This subcommunicator would be involved in all the tasks to be performed in order to solve the flow problem in parallel. For example, with reference to Figure 3.1, let us assume that we decide to work with 15 processes. The default MPI communicator would be

$$mpi\_comm\_world = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14\}$$

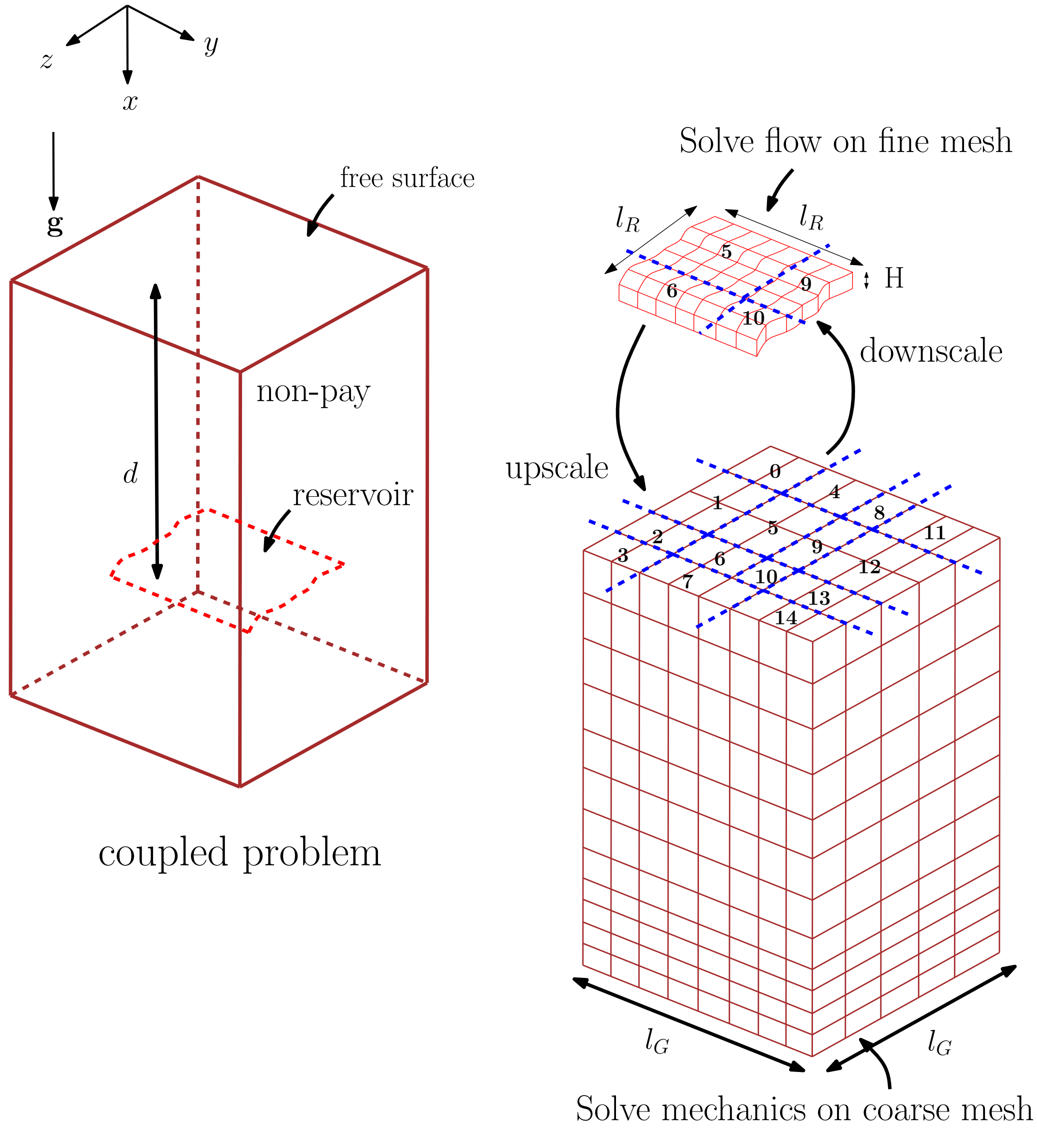

Figure 3.1: The MPI based parallelism involved in IPARS. The numbers are process ranks, and the division of elements among processes is areal. For this particular case, the default MPI communicator has 15 processes. Since the flow domain is a subset of the geomechanical domain, not all processes would occupy flow elements. The MPI subcommunicator would have only those processes that occupy flow elements. Only the processes involved in the subcommunicator would be involved in the construction of grid-to-grid projection operators.

Thereafter, based on the actual location of the areal $(y - z)$ boundaries of the flow domain, the flow elements are assigned to processes based on the areal boundaries of the processes. With reference to Figure 3.1, processes with ranks $5, 6, 9, 10$ occupy elements of the flow domain. Thus, the subcommunicator would be

$$flow\_comm = \{5, 6, 9, 10\}$$

This means that processes with ranks $0, 1, 2, 3, 4, 7, 8, 11, 12, 13, 14$ do not occupy flow elements, and need not participate in construction of grid-to-grid projection operators. The non-participation of as many as 11 processes in the construction of projection operators, and indeed the flow solve, substantially reduces the memory load on the supercomputer. Data structures (with concomitant memory allocation) required for the flow solve need only be assigned to 4 processes instead of the usual 15 for this particular case.

3.2 Identifying pairs that intersect with one another

The global upscaling and downscaling operators are constructed in the pre-processing step by applying a local mapping procedure to each pair of overlapping finite elements E^p and E^f of the two grids. As shown in Figure 3.2, we denote the intersection polyhedron of E^p and E^f by $\mathcal{E}$, and further denote the partitions of E^p and E^f by $\mathcal{I}^{E^p}$ and $\mathcal{I}^{E^f}$ respectively such that

$$E^p = \bigcup_{\mathcal{E} \in \mathcal{I}^{E^p}} \mathcal{E} \qquad \text{where} \qquad \mathcal{I}^{E^p} \equiv \left\{ \mathcal{E} : \mathcal{E} \equiv E^p \cap E^f \ \forall E^f \in \mathcal{T}_h^f \right\}$$

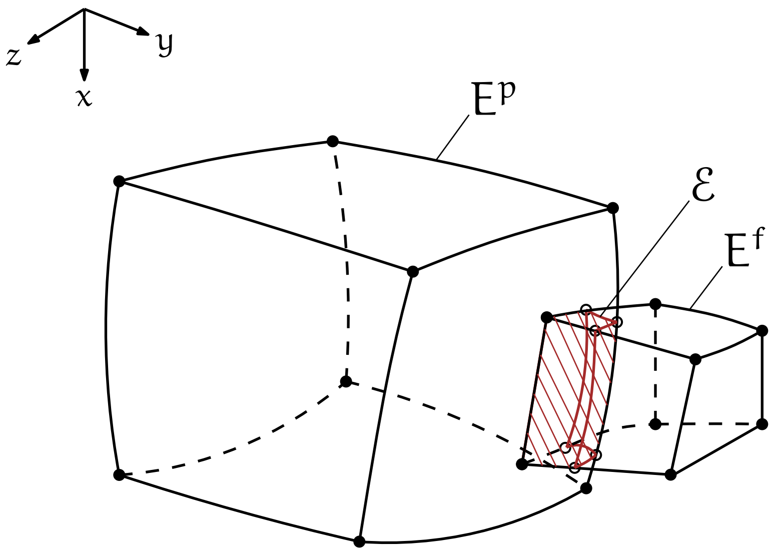

Figure 3.2: Construction of global up and down scaling operators are designed around the assembly of local up and down scaling operators.

$$E^f = \bigcup_{\mathcal{E}\in\mathcal{T}^{E^f}} \mathcal{E} \qquad \text{where} \qquad \mathcal{T}^{E^f} = \left\{\mathcal{E} : \mathcal{E} = E^f \cap E^p \ \forall E^p \in \mathcal{T}^p_H\right\}$$

We first need to identify pairs of flow and poromechanics elements that intersect one another. Each pair will have one flow element and one poromechanics element. To determine whether a flow element is geometrically disjoint to a poromechanics element, we need to design an algorithm with all 8 vertices of the flow element. If all vertices of the flow element are inside the poromechanics element, then the flow element is inside the poromechanics element. If all vertices of the flow element are outside the poromechanics element, then the flow element is geometrically disjoint to the poromechanics element. If neither of the above two cases is true, then the flow element intersects the poromechanics element. A graphic description of the algorithm to determine if a point is inside or outside a finite element is given in Figure 3.3. The algorithm itself is given in Algorithm 2.

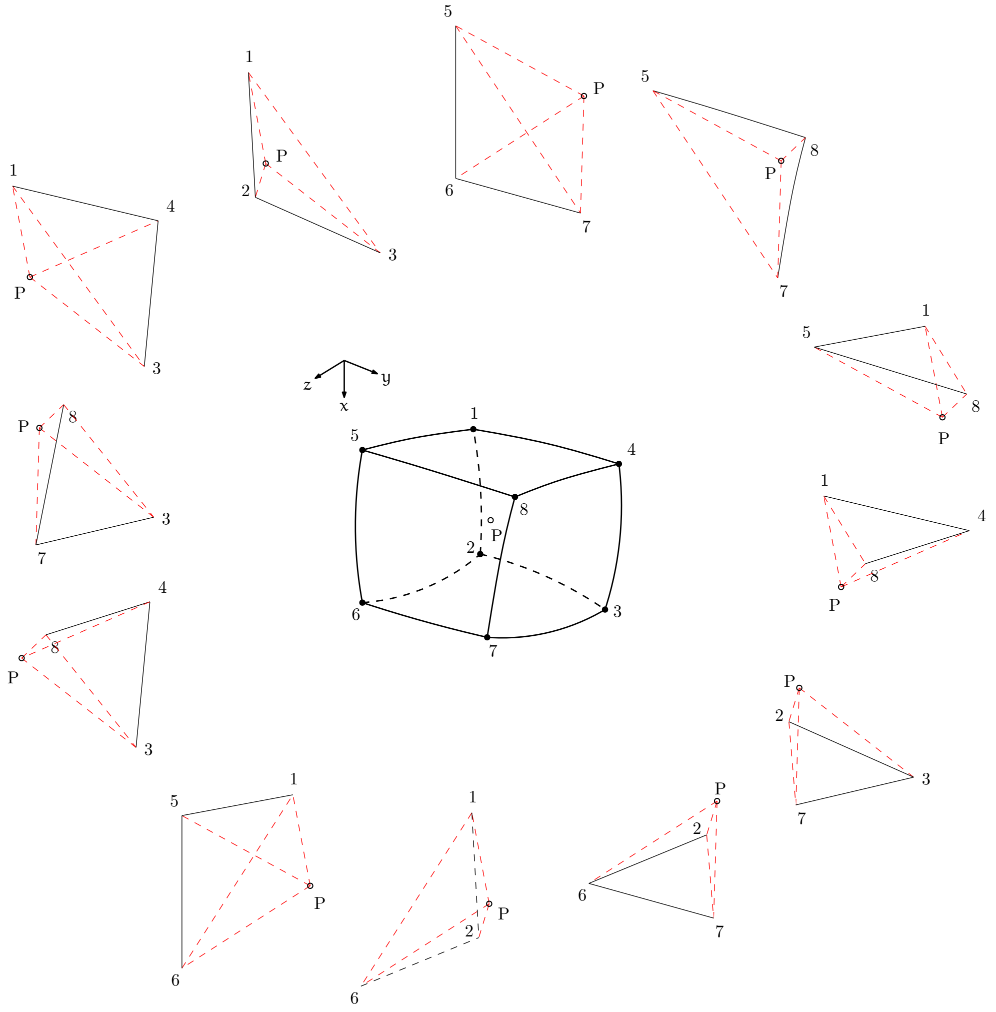

Figure 3.3: Determing whether point P is inside or outside the finite element 1-2-3-4-5-6-7-8. If point P is inside the finite element, the sum of volumes of the 12 tetrahedra formed by P with the vertices of the finite element would be roughly equal to volume of the finite element. If point P is outside the finite element, the sum of volumes of the 12 tetrahedra would be greater than the volume of the finite element

Algorithm 2 Determining whether flow element is inside, or outside, or intersects a poromechanics element

```
facemap= $\begin{bmatrix} 1 & 5 & 8 \\ 1 & 8 & 4 \\ 2 & 3 & 7 \\ 2 & 7 & 6 \\ 1 & 2 & 6 \\ 1 & 6 & 5 \\ 3 & 4 & 8 \\ 3 & 8 & 7 \\ 1 & 4 & 3 \\ 1 & 3 & 2 \\ 5 & 6 & 7 \\ 5 & 7 & 8 \end{bmatrix}$
porovol ← volume of poromechanics element
outcount ← 0
for i=1 to 8 do
    volsum ← 0
    Point P is the $i^{th}$ vertex of flow element
    for j=1 to 12 do
        Tet ← [P,facemap(j,:)]        ▷ Tet is the tetrahedron with point P, and the three numbered vertices corresponding to facemap(j,:) of the poromechanics element as the four vertices
        Tetvol ← volume of Tet
        volsum ← volsum + Tetvol
    if volsum > porovol then
        outcount ← outcount + 1
if outcount=0 then
    flow element is inside poromechanics element
else if outcount=8 then
    flow element is outside poromechanics element
else
    flow element intersects poromechanics element
```

3.3 Upscaling pore pressure and bulk density for each pair

Let p^{E^f} and ρ^{E^f} represent the piecewise constant pore pressure and fluid density respectively at E^f such that

$$p^{\mathcal{E}} = p^{E^f}, \qquad \rho^{\mathcal{E}} = \rho^{E^f} \qquad\qquad (\text{if } \mathcal{E} \in E^f)$$

The pore pressures are upscaled via the force vector term $\int\limits_{E^p} \mathbf{B}^T \alpha p \mathbf{I}$ in (2.37) as follows

$$\int\limits_{E^p} \mathbf{B}^T \alpha p \mathbf{I} = \int\limits_{E^p} \mathbf{B}^T \alpha \bigg(\sum_{\mathcal{E} \in \mathcal{I}^{E^p}} \frac{Meas(\mathcal{E})}{Meas(E^p)} p^{\mathcal{E}} \bigg) \mathbf{I} \tag{3.1}$$

The bulk densities are upscaled via the force vector term $\int\limits_{E^p} \mathbf{N}^T \rho \phi \mathbf{g}$ in (2.37) as follows

$$\int\limits_{E^p} \mathbf{N}^T \rho \phi \mathbf{g} = \int\limits_{E^p} \mathbf{N}^T \bigg(\sum_{\mathcal{E} \in \mathcal{I}^{E^p}} \frac{Meas(\mathcal{E})}{Meas(E^p)} \rho^{\mathcal{E}} \phi^{\mathcal{E}} \bigg) \mathbf{g} \tag{3.2}$$

In essence, the upscaled pore pressures and bulk densities on E^p are local volume averages over $\mathcal{I}^{E^p}$ of the information obtained after the flow solve. The reason for choosing the local volume average approach for upscaling of the fine scale pore pressure and product of fine scale fluid density and porosity is the following: Suppose v is any variable being projected onto a space of constants $W_H \equiv \mathbb{P}_0(\mathscr{T}_H^p)$ defined on $\mathscr{T}_H^p$. Then, the Galerkin orthogonal projection of v onto W_H would be

$$\int\limits_{\Omega^p} (v - \mathcal{R}v) w = 0 \qquad \forall\ w \in W_H \tag{3.3}$$

where $\mathcal{R}v \in W_H$ is the projected quantity. Let v' denote the restriction of v to E^p and let w' denote the restriction of w to E^p. We rewrite (3.3) as

$$\sum_{E^p \in \mathcal{T}_H^p} \int\limits_{E^p} (v' - \mathcal{R}v) w' = 0$$

which, in lieu of the linear independence of the basis w' of the space W_H, implies that

$$\mathcal{R}v = \frac{1}{Meas(E^p)} \int\limits_{E^p} v' \qquad (\forall\ E^p \in \mathscr{T}_H^p) \tag{3.4}$$

If v' is defined by the discontinuous piecewise constants $v^{\mathcal{E}}$ on $\mathcal{J}^{E^p}$ $(\forall\ E^p \in \mathscr{T}_H^p)$, we get

$$\int\limits_{E^p} v' \equiv \sum_{\mathcal{E} \in \mathcal{J}^{E^p}} \int\limits_{\mathcal{E}} v^{\mathcal{E}} = \sum_{\mathcal{E} \in \mathcal{J}^{E^p}} v^{\mathcal{E}} Meas(\mathcal{E}) \qquad (\forall\ E^p \in \mathscr{T}_H^p) \tag{3.5}$$

In lieu of (3.4) and (3.5), we get

$$\mathcal{R}v = \frac{1}{Meas(E^p)} \sum_{\mathcal{E} \in \mathcal{J}^{E^p}} v^{\mathcal{E}} Meas(\mathcal{E}) \equiv \sum_{\mathcal{E} \in \mathcal{J}^{E^p}} \frac{Meas(\mathcal{E})}{Meas(E^p)} v^{\mathcal{E}} \qquad (\forall\ E^p \in \mathscr{T}_H^p)$$

3.4 Downscaling Lagrangian porosity for each pair

Let $W_h \equiv \mathbb{P}_0(\mathscr{T}_h^f)$ represent the space of constants defined on $\mathscr{T}_h^f$. Let $\mathcal{P}\phi^* \in W_h$ be the prolongation of the coarse scale Lagrangian porosity $\phi^*(\mathbf{x}^p)$, $\mathbf{x}^p \in \Omega^p \cap \Omega^f$ onto W_h. Define $\mathcal{P}$ by

$$\int\limits_{\Omega^f} \big(\phi^*(\mathbf{x}) - (\mathcal{P}\phi^*)(\mathbf{x})\big)\, w = 0 \qquad \forall w \in W_h \tag{3.6}$$

Let $\phi'(\mathbf{x}')$, $\mathbf{x}' \in E^f$ denote the restriction of $\phi^*(\mathbf{x})$ to E^f and let w' denote the restriction of w to E^f. We rewrite (3.6) as

$$\sum_{E^f \in \mathcal{T}_h^f} \int\limits_{E^f} (\phi'(\mathbf{x}') - \mathcal{P}\phi^*) w' = 0 \tag{3.7}$$

which, in lieu of the linear independence of the basis w' of the space W_h, implies that

$$\mathcal{P}\phi^* = \frac{1}{Meas(E^f)} \int\limits_{E^f} \phi'(\mathbf{x}') \qquad (\forall\ E^f \in \mathscr{T}_h^f)$$

Let $\phi^{*^{E^p}}$ be the piecewise constant Lagrangian porosity at E^p $(\forall\ E^p \in \mathscr{T}_H^p)$ such that

$$\phi^{*^{\mathcal{E}}} = \phi^{*^{E^p}} \qquad (\text{if } \mathcal{E} \in E^p)$$

Then $\phi'(\mathbf{x}')$ is defined by discontinuous piecewise constants over $\mathcal{J}^{E^f}$ as

$$\phi'(\mathbf{x}') = \phi^{*^{\mathcal{E}}} = \phi^{*^{E^p}} \qquad (\forall \mathbf{x}' \in \mathcal{E} \text{ and } \mathcal{E} \in E^p)$$

In lieu of the above, we get

$$\mathcal{P}\phi^* = \frac{\sum\limits_{\mathcal{E} \in \mathcal{J}^{E^f}} \phi^{*^{E^p}} Meas(\mathcal{E})}{Meas(E^f)} \qquad (\mathcal{E} = E^f \cap E^p)$$

3.5 Local operators : implementation

- The first step is to obtain the equations of the faces of E^p and E^f as explained in Section 3.5.1.

- The next step is to design an algorithm that uses the equations of the element faces to obtain points on the periphery of the intersection polyhedron as explained in Section 3.5.2.
- The final step is to use the set of points obtained on the periphery of the intersection polyhedron to determine its measure as explained in Section 3.5.3.

3.5.1 Obtaining equations of the element faces

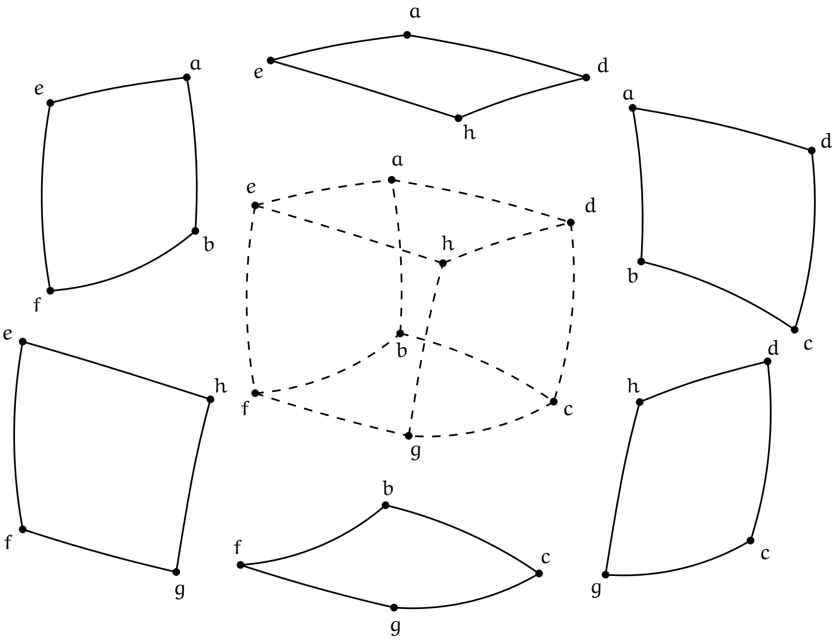

Figure 3.4: A representation of hexahedral element $E \equiv abcdefgh$ with its six faces $aehd$, $abfe$, $ehgf$, $bcgf$, $cdhg$ and $adcb$. The coordinate information of the four vertices of each of the faces is used to obtain its equation.

As the procedure invoked in obtaining the equations of the element faces applies to both E^f and E^p, we use the notation E for hexahedral element without regard to whether it is a flow element E^f or poromechanics element E^p. Let $\mathcal{S}(\mathbf{x}) = 0$, $\mathbf{x} \equiv (x, y, z) \in e$ be the equation of face e of element E with its vertices $\mathbf{v}_i \equiv (x_i, y_i, z_i)$, $i = 1, 2, 3, 4$. A representation of E with its faces is provided in Figure 3.4. Define $\mathcal{S}(\mathbf{x})$ by a trilinear as

$$\mathcal{S}(\mathbf{x}) = \left\{ \begin{array}{cccccccc} xyz & xy & yz & xz & x & y & z & 1 \end{array} \right\} \mathbf{c}_{8\times 1} \tag{3.8}$$

where $\mathbf{c}_{8\times 1}$ is the vector of coefficients to be determined. Since $\mathcal{S}(\mathbf{x}) = 0$ is satisfied at each of the four vertices defining the face, we get the system of equations for $\mathbf{c}$

$$\overbrace{\begin{bmatrix} x_1y_1z_1 & x_1y_1 & y_1z_1 & x_1z_1 & x_1 & y_1 & z_1 & 1 \\ x_2y_2z_2 & x_2y_2 & y_2z_2 & x_2z_2 & x_2 & y_2 & z_2 & 1 \\ x_3y_3z_3 & x_3y_3 & y_3z_3 & x_3z_3 & x_3 & y_3 & z_3 & 1 \\ x_4y_4z_4 & x_4y_4 & y_4z_4 & x_4z_4 & x_4 & y_4 & z_4 & 1 \end{bmatrix}}^{\mathbf{M}_{4\times 8}} \mathbf{c}_{8\times 1} = \left\{ \begin{array}{c} 0 \\ 0 \\ 0 \\ 0 \end{array} \right\}_{4\times 1}$$

It is clear that $\mathbf{c}$ belongs to the nullspace of $\mathbf{M}$. With that in mind, we get the SVD of $\mathbf{M}$ as

$$\mathbf{M}_{4\times 8} = \mathbf{U}_{4\times 4} \boldsymbol{\sigma}_{4\times 8} \mathbf{V}^T_{8\times 8} \tag{3.9}$$

where $\boldsymbol{\sigma} = diag(\sigma_1, .., \sigma_r)$ is diagonal matrix of singular values of $\mathbf{M}$ and the columns of $\mathbf{U}$ and $\mathbf{V}$ are left and right singular vectors of $\mathbf{M}$ respectively. Since the nullspace of $\mathbf{M}$ is spanned by right singular vectors corresponding to the vanishing singular values of $\mathbf{M}$, we express $\mathbf{c}$ as

$$\mathbf{c}_{8\times 1} = \begin{bmatrix} \mathbf{V}[:, r+1] & . & . & \mathbf{V}[:, 8] \end{bmatrix}_{8\times(8-r)} \boldsymbol{\kappa}_{(8-r)\times 1} \tag{3.10}$$

where $\boldsymbol{\kappa}$ is the vector of coefficients and r is rank of $\mathbf{M}$. The objective now is to determine $\boldsymbol{\kappa}$. First, using (3.8), we obtain an expression for the gradient $\nabla\mathcal{S}(\mathbf{x})$ of $\mathcal{S}(\mathbf{x})$ as

$$\nabla\mathcal{S}(\mathbf{x}) = \overbrace{\begin{bmatrix} yz & y & 0 & z & 1 & 0 & 0 & 0 \\ xz & x & z & 0 & 0 & 1 & 0 & 0 \\ xy & 0 & y & x & 0 & 0 & 1 & 0 \end{bmatrix}}^{\mathbf{H}(x,y,z)_{3\times 8}} \begin{bmatrix} \mathbf{V}[:, r+1] & . & . & \mathbf{V}[:, 8] \end{bmatrix}_{8\times(8-r)} \boldsymbol{\kappa}_{(8-r)\times 1} \tag{3.11}$$

Let $\hat{\mathcal{S}}(\hat{\mathbf{x}})$ be corresponding definition on face $\hat{e}$ of reference element $\hat{E}$ of $\mathcal{S}(\mathbf{x})$ on face e of actual element E. Then, from (2.15) and (2.16),

$$\nabla\mathcal{S}(\mathbf{x}) = (DF_E)^{-T}(\hat{\mathbf{x}})\,\hat{\nabla}\hat{\mathcal{S}}(\hat{e}) \tag{3.12}$$

where $\hat{\nabla}\hat{\mathcal{S}}(\hat{e})$ can be either $\begin{bmatrix}1 & 0 & 0\end{bmatrix}^T$, $\begin{bmatrix}0 & 1 & 0\end{bmatrix}^T$ or $\begin{bmatrix}0 & 0 & 1\end{bmatrix}^T$ depending on whether $\hat{e}$ is normal to $\hat{x}$, $\hat{y}$ or $\hat{z}$ axis. Equating (6.10) and (3.12) for all four vertices of $e \in E$, we get the following system of equations for $\boldsymbol{\kappa}_{(8-r)\times 1}$

$$\begin{bmatrix} \mathbf{H}(x_1, y_1, z_1) \\ \mathbf{H}(x_2, y_2, z_2) \\ \mathbf{H}(x_3, y_3, z_3) \\ \mathbf{H}(x_4, y_4, z_4) \end{bmatrix}_{12\times 8} \begin{bmatrix} \mathbf{V}[:, r+1] & . & . & \mathbf{V}[:, 8] \end{bmatrix}_{8\times(8-r)} \boldsymbol{\kappa}_{(8-r)\times 1} = \mathbf{B}_{12\times 1} \tag{3.13}$$

where $\mathbf{B}$ is obtained as

$$\mathbf{B}[(i-1)*3+1 \rightarrow i*3, 1] = (DF_E)^{-T}(\hat{\mathbf{v}}_{\mathbf{i}})\,\hat{\nabla}\hat{\mathcal{S}}(\hat{e})$$

where $\hat{\mathbf{v}}_{\mathbf{i}}$, $i = 1, 2, 3, 4$ on $\hat{e} \in \hat{E}$ is the corresponding definition of $\mathbf{v}_i$, $i = 1, 2, 3, 4$ on $e \in E$. The solution $\boldsymbol{\kappa}$ of (3.13) is substituted into (3.10) to obtain $\mathbf{c}$, which is then substituted into (3.8) to obtain the polynomial expression of $\mathcal{S}(\mathbf{x})$.

3.5.2 Obtaining points on the periphery of intersection polyhedron using surface-surface intersections

Algorithm 3 obtains points on the periphery of $\mathcal{E}$ by tracing the intersections of the faces of E^f and E^p. The starting points for the curve traces are the set of vertices of E^f inside E^p.

To understand the construct of the algorithm, we refer to a case depicted in 3.5. Red arrows represent the set $\mathcal{Q}^{ff}$ i.e. traces of the intersections of the faces of E^f. Blue arrows represent the set $\mathcal{Q}^{fp}$ i.e. traces of the intersections of the faces of E^f with the faces of E^p. Green arrows represent the set $\mathcal{Q}^{pp}$ i.e. traces of the intersections of the faces of E^p. The nested loop structure in Algorithm 3 traces out elements of the set $\mathcal{Q}_{ff}$, $\mathcal{Q}_{fp}$ and $\mathcal{Q}_{pp}$ sequentially. Representing $\mathcal{S}^f_{x-}$, $\mathcal{S}^f_{x+}$, $\mathcal{S}^f_{y-}$, $\mathcal{S}^f_{y+}$, $\mathcal{S}^f_{z-}$ and $\mathcal{S}^f_{z+}$ as the faces of the flow element normal to the $x-$, $x+$, $y-$, $y+$, $z-$ and $z+$ axes respectively and $\mathcal{S}^p_{x-}$, $\mathcal{S}^p_{x+}$, $\mathcal{S}^p_{y-}$, $\mathcal{S}^p_{y+}$, $\mathcal{S}^p_{z-}$ and $\mathcal{S}^p_{z+}$ as the faces of the poromechanics element normal to the $x-$, $x+$, $y-$, $y+$, $z-$ and $z+$ axes respectively, the curve traces are

$$AD \equiv \mathcal{S}^f_{x-} \cap \mathcal{S}^f_{z+} \rightarrow DH \equiv \mathcal{S}^f_{z+} \cap \mathcal{S}^p_{y+} \qquad \text{(Sequence 1, Starting point } A \text{ on } \mathcal{S}^f_{x-}\text{)}$$

$$EF \equiv \mathcal{S}^f_{x+} \cap \mathcal{S}^f_{y-} \rightarrow FB \equiv \mathcal{S}^f_{y-} \cap \mathcal{S}^p_{z-} \qquad \text{(Sequence 2, Starting point } E \text{ on } \mathcal{S}^f_{x+}\text{)}$$

$$AB \equiv \mathcal{S}^f_{y-} \cap \mathcal{S}^f_{x-} \rightarrow BC \equiv \mathcal{S}^f_{x-} \cap \mathcal{S}^p_{z-} \rightarrow CG \equiv \mathcal{S}^p_{z-} \cap \mathcal{S}^p_{y+} \qquad \text{(Sequence 3, Starting point } A \text{ on } \mathcal{S}^f_{y-}\text{)}$$

$$EA \equiv \mathcal{S}^f_{y-} \cap \mathcal{S}^f_{z-} \qquad \text{(Sequence 4, Starting point } E \text{ on } \mathcal{S}^f_{z-}\text{)}$$

$$EH \equiv \mathcal{S}^f_{z+} \cap \mathcal{S}^f_{x+} \rightarrow HG \equiv \mathcal{S}^f_{x+} \cap \mathcal{S}^p_{y+} \rightarrow GC \equiv \mathcal{S}^p_{y+} \cap \mathcal{S}^p_{z-} \qquad \text{(Sequence 5, Starting point } E \text{ on } \mathcal{S}^f_{z+}\text{)}$$

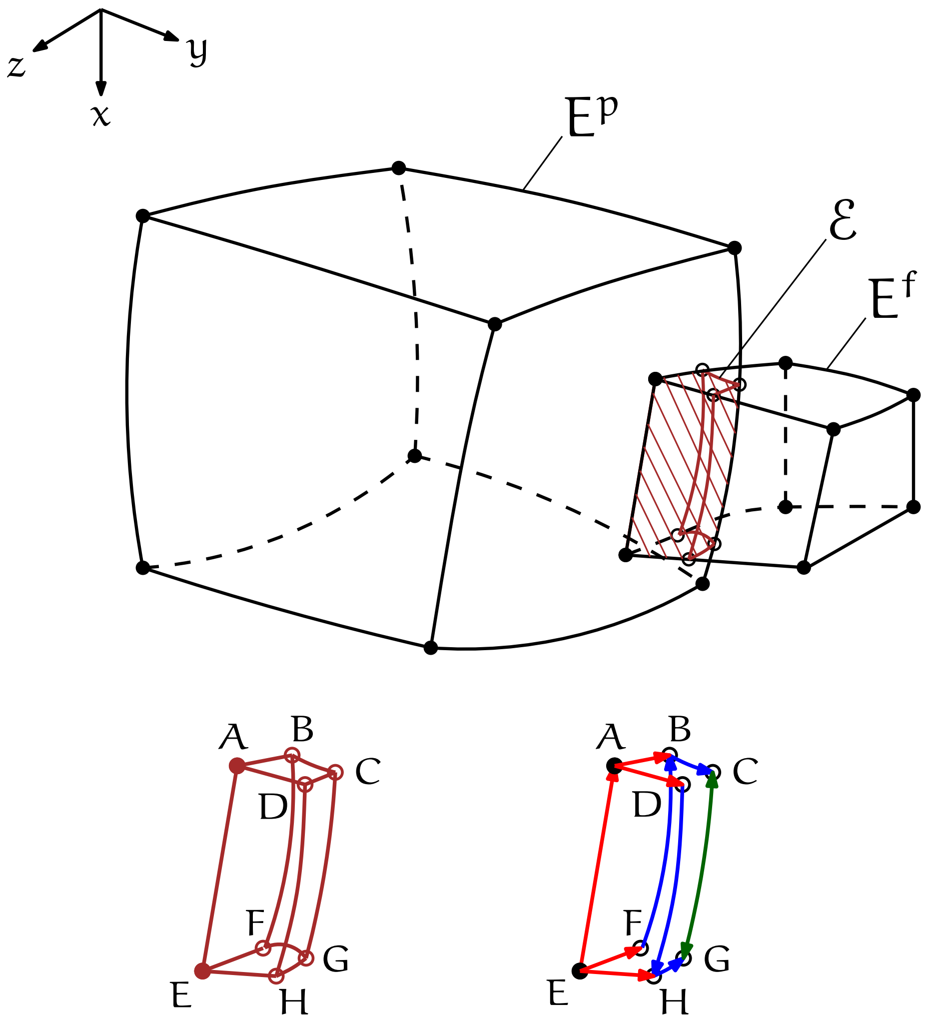

Figure 3.5: Solid circles representing points A and E are the vertices of E^f inside E^p. Hollow circles representing points B, C, D, E, F, G and H are the end points of the curve traces obtained using the surface-surface intersections algorithm. The arrows represent the direction of the curve traces. It is important to note that Figure 3.5 is only a depiction of $\mathcal{E}$ and that there is no restriction whatsoever that it be 8-noded.

3.5.2.1 The algorithm

Algorithm 3 Obtaining points on the periphery of intersection polyhedron

```
if $E^f \cap E^p \leftarrow \emptyset$ then
    $\mathcal{E} \leftarrow \emptyset$    ▷ If $E^f$ and $E^p$ are geometrically disjoint, $\mathcal{E}$ is a null set
else if $E^f \subset E^p$ then
    $\mathcal{E} \leftarrow E^f$    ▷ If $E^f$ is inside $E^p$, $\mathcal{E} \equiv E^f$
else    ▷ If $E^f$ and $E^p$ intersect
    $\mathcal{N} \leftarrow \emptyset$    ▷ Initialize the set of points $\mathcal{N}$ on periphery of $\mathcal{E}$
    for $i = 1, .., 6$ do    ▷ Loop over six faces of $E^f$
        $\mathbf{q} \in \mathcal{A} \cap \mathcal{S}_i^f$    ▷ $\mathcal{A}$ is set of vertices of $E^f$ inside $E^p$, $\mathcal{S}_i^f$ is the $i^{th}$ face of $E^f$, $\mathbf{q}$ is the starting point for the curve trace
        for $j = 1, .., 6$ do    ▷ Loop over six faces of $E^f$
            while $\mathbf{q} \in E^f \wedge \mathbf{q} \in E^p$ do
                $\mathbf{q} \leftarrow$ TRACE $(\mathbf{q}, \mathcal{S}_i^f, \mathcal{S}_j^f)$    ▷ $\mathcal{S}_j^f$ is the $j^{th}$ face of $E^f$
            $\mathcal{Q}^{ff} \leftarrow \mathbf{q}$    ▷ Final curve trace on the intersection of $\mathcal{S}_i^f$ with $\mathcal{S}_j^f$
            for $k = 1, .., 6$ do    ▷ Loop over six faces of $E^p$
                while $\mathbf{q} \in E^f \wedge \mathbf{q} \in E^p$ do
                    $\mathbf{q} \leftarrow$ TRACE $(\mathbf{q}, \mathcal{S}_j^f, \mathcal{S}_k^p)$    ▷ $\mathcal{S}_k^p$ is the $k^{th}$ face of $E^p$
                $\mathcal{Q}^{fp} \leftarrow \mathbf{q}$    ▷ Final curve trace on the intersection of $\mathcal{S}_j^f$ with $\mathcal{S}_k^p$
                for $l = 1, .., 6$ do    ▷ Loop over six faces of $E^p$
                    while $\mathbf{q} \in E^f \wedge \mathbf{q} \in E^p$ do
                        $\mathbf{q} \leftarrow$ TRACE $(\mathbf{q}, \mathcal{S}_k^p, \mathcal{S}_l^p)$    ▷ $\mathcal{S}_l^p$ is the $l^{th}$ face of $E^p$
                    $\mathcal{Q}^{pp} \leftarrow \mathbf{q}$    ▷ Final curve trace on the intersection of $\mathcal{S}_k^p$ with $\mathcal{S}_l^p$
        $\mathcal{N} \leftarrow \mathcal{N} \cup (\mathcal{Q}^{ff} \oplus \mathcal{Q}^{fp} \oplus \mathcal{Q}^{pp})$    ▷ Union of the curve traces
    $\mathcal{V}_{\mathcal{E}} \leftarrow \mathcal{A} \oplus \mathcal{N}$    ▷ Union of $\mathcal{A}$ and $\mathcal{N}$
```

Curve tracing : the function $TRACE(\mathbf{q}, \mathcal{S}_1, \mathcal{S}_2)$

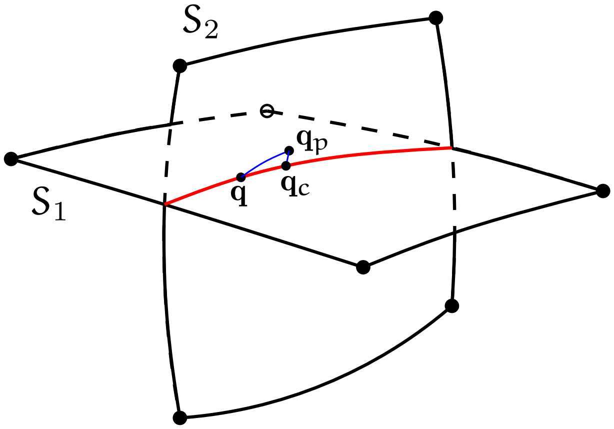

Figure 3.6: $\mathcal{S}_1 \cap \mathcal{S}_2$ represented by red solid line. $\mathbf{q}_p$ is the predictor to the trace of $\mathcal{S}_1 \cap \mathcal{S}_2$. $\mathbf{q}_c$ is the corrector to $\mathbf{q}_p$.

The curve tracing is based on the predictor-corrector approach of Bajaj et al. [5]. The predictor gives a second order Taylor approximant to the trace of the intersection curve and the corrector refines the approximant to a point on the intersection curve using the Newton method. With reference to Figure 3.6, a predictor to the trace of the intersection of surfaces $\mathcal{S}_1$ and $\mathcal{S}_2$ is stored in $\mathbf{q}_p$ and the corrector to $\mathbf{q}_p$ is stored in $\mathbf{q}_c$.

Obtaining $\mathbf{q}_p$

The intersection curve $\mathbf{r}$ of $\mathcal{S}_1$ and $\mathcal{S}_2$ with initial point $\mathbf{q}$ is expressed as

$$\mathbf{r}(s) = \mathbf{q} + s\mathbf{r}'(\mathbf{q}) + \frac{s^2}{2!}\mathbf{r}''(\mathbf{q}) + \mathbf{e}(s) = \mathbf{q}_p(s) + \mathbf{e}(s)$$

where s is an arc length parameter, $\mathbf{r}'(\mathbf{q})$ is unit tangent to curve at $\mathbf{q}$, $\mathbf{r}''(\mathbf{q})$ is curvature at $\mathbf{q}$ and $\mathbf{e}(s) = O(s^3)$ is the error. We assume $s = 0.1$, a value small enough to make $\mathbf{q}_p(s)$ an accurate estimate of $\mathbf{r}(s)$ i.e. $|\mathbf{e}(s)| << |\mathbf{p}(s)|$. As long as $\mathbf{q}$ is not singular on $\mathcal{S}_1$ or on $\mathcal{S}_2$, the surface gradients $\nabla\mathcal{S}_1(\mathbf{q})$ and $\mathcal{S}_2(\mathbf{q})$ are linearly independent (see Bajaj et al. [5]) and the unit tangent vector $\mathbf{r}'(\mathbf{q})$ is obtained as

$$\mathbf{r}'(\mathbf{q}) = \frac{\nabla\mathcal{S}_1(\mathbf{q}) \times \nabla\mathcal{S}_2(\mathbf{q})}{\|\nabla\mathcal{S}_1(\mathbf{q}) \times \nabla\mathcal{S}_2(\mathbf{q})\|} \tag{3.14}$$

It follows that $\mathbf{r}'(\mathbf{q})$ is perpendicular to both the surface gradients such that

$$\nabla\mathcal{S}_1(\mathbf{q}) \cdot \mathbf{r}'(\mathbf{q}) = \nabla\mathcal{S}_2(\mathbf{q}) \cdot \mathbf{r}'(\mathbf{q}) = 0$$

implying the vectors $\nabla\mathcal{S}_1(\mathbf{q})$, $\nabla\mathcal{S}_2(\mathbf{q})$ and $\mathbf{r}'(\mathbf{q})$ are linearly independent. We express $\mathbf{r}''(\mathbf{q})$ as a linear combination of $\nabla\mathcal{S}_1(\mathbf{q})$, $\nabla\mathcal{S}_2(\mathbf{q})$ and $\mathbf{r}'(\mathbf{q})$ as

$$\mathbf{r}''(\mathbf{q}) = \alpha\mathbf{r}'(\mathbf{q}) + \beta\nabla\mathcal{S}_1(\mathbf{q}) + \gamma\nabla\mathcal{S}_2(\mathbf{q}) \tag{3.15}$$

The points on $\mathbf{r}(s)$ are solutions of $\mathcal{S}_1(\mathbf{r}(s)) = \mathcal{S}_1(\mathbf{r}(s)) = 0$. The Taylor expansion of $\mathcal{S}_j(\mathbf{r}(s))$, $j = 1, 2$ with $\mathbf{q} \equiv \mathbf{r}'(0)$ is

$$\mathcal{S}_j(\mathbf{r}(s)) = \mathcal{S}_j(\mathbf{q}) + s\nabla\mathcal{S}_j(\mathbf{q}) \cdot \mathbf{r}'(\mathbf{q}) + \frac{s^2}{2!}(\nabla\mathcal{S}_j(\mathbf{q}) \cdot \mathbf{r}''(\mathbf{q}) + \mathbf{r}'(\mathbf{q}) \cdot H_{\mathcal{S}_j}(\mathbf{q}) \cdot \mathbf{r}'(\mathbf{q}))$$

where $H_{\mathcal{S}_j}(\mathbf{q})$ is the Hessian of the surface $\mathcal{S}_j$ evaluated at $\mathbf{q}$ as follows

$$H_{\mathcal{S}_j}(\mathbf{q}) = \begin{bmatrix} \frac{\partial^2\mathcal{S}_j}{\partial x^2} & \frac{\partial^2\mathcal{S}_j}{\partial x\partial y} & \frac{\partial^2\mathcal{S}_j}{\partial x\partial z} \\ \frac{\partial^2\mathcal{S}_j}{\partial y\partial x} & \frac{\partial^2\mathcal{S}_j}{\partial y^2} & \frac{\partial^2\mathcal{S}_j}{\partial y\partial z} \\ \frac{\partial^2\mathcal{S}_j}{\partial z\partial x} & \frac{\partial^2\mathcal{S}_j}{\partial z\partial y} & \frac{\partial^2\mathcal{S}_j}{\partial z^2} \end{bmatrix}_{3\times 3}$$

where the diagonal entries of $H_{\mathcal{S}_j}(\mathbf{q})$ are zero and the off diagonal entries are

$$\frac{\partial^2\mathcal{S}}{\partial x\partial y} \equiv \frac{\partial^2\mathcal{S}}{\partial y\partial x} = \left\{ \begin{matrix} z & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{matrix} \right\} \mathbf{c}$$

$$\frac{\partial^2 \mathcal{S}}{\partial x \partial z} \equiv \frac{\partial^2 \mathcal{S}}{\partial z \partial x} = \left\{ \begin{array}{cccccccc} y & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{array} \right\} \mathbf{c}$$

$$\frac{\partial^2 \mathcal{S}}{\partial y \partial z} \equiv \frac{\partial^2 \mathcal{S}}{\partial z \partial y} = \left\{ \begin{array}{cccccccc} x & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{array} \right\} \mathbf{c}$$

where x, y and z are the coordinates of the point $\mathbf{q}$ and $\mathbf{c}$ is the vector of coefficients corresponding to the definition of $\mathcal{S}_j$.

Since $\mathcal{S}_j(\mathbf{r}(s)) \equiv 0$, $j = 1, 2$, the coefficient of each power of s in $\mathcal{S}_j(\mathbf{r}(s))$ must be zero. We already know that the coefficient of s in $\mathcal{S}_j(\mathbf{r}(s))$ is zero i.e. $\nabla \mathcal{S}_j(\mathbf{q}) \cdot \mathbf{r}'(\mathbf{q}) = 0$, $j = 1, 2$. Equating the coefficient of s^2 in $\mathcal{S}_j(\mathbf{r}(s))$ to zero, we get

$$\nabla \mathcal{S}_j(\mathbf{q}) \cdot \mathbf{r}''(\mathbf{q}) = -\mathbf{r}'(\mathbf{q}) \cdot H_{\mathcal{S}_j}(\mathbf{q}) \cdot \mathbf{r}'(\mathbf{q})$$

$$\nabla \mathcal{S}_j(\mathbf{q}) \cdot (\alpha \mathbf{r}'(\mathbf{q}) + \beta \nabla \mathcal{S}_1(\mathbf{q}) + \gamma \nabla \mathcal{S}_2(\mathbf{q})) = -\mathbf{r}'(\mathbf{q}) \cdot H_{\mathcal{S}_j}(\mathbf{q}) \cdot \mathbf{r}'(\mathbf{q})$$

and noting again that $\nabla \mathcal{S}_j(\mathbf{q}) \cdot \mathbf{r}'(\mathbf{q}) = 0$, $j = 1, 2$, we get the following system of equations for β and γ

$$\begin{bmatrix} \nabla \mathcal{S}_1(\mathbf{q}) \cdot \nabla \mathcal{S}_1(\mathbf{q}) & \nabla \mathcal{S}_1(\mathbf{q}) \cdot \nabla \mathcal{S}_2(\mathbf{q}) \\ \nabla \mathcal{S}_2(\mathbf{q}) \cdot \nabla \mathcal{S}_1(\mathbf{q}) & \nabla \mathcal{S}_2(\mathbf{q}) \cdot \nabla \mathcal{S}_2(\mathbf{q}) \end{bmatrix} \left\{ \begin{array}{c} \beta \\ \gamma \end{array} \right\} = - \left\{ \begin{array}{c} \mathbf{r}'(\mathbf{q}) \cdot H_{\mathcal{S}_1}(\mathbf{q}) \cdot \mathbf{r}'(\mathbf{q}) \\ \mathbf{r}'(\mathbf{q}) \cdot H_{\mathcal{S}_2}(\mathbf{q}) \cdot \mathbf{r}'(\mathbf{q}) \end{array} \right\} \tag{3.16}$$

Solution of (3.16) and the choice $\alpha = 0$ leads to a unique vector $\mathbf{r}''(\mathbf{q})$ in (3.15). The second order interpolant $\mathbf{q}_p(s)$ is finally obtained as

$$\mathbf{q}_p(s) = \mathbf{q} + 0.1 \frac{\nabla \mathcal{S}_1(\mathbf{q}) \times \nabla \mathcal{S}_2(\mathbf{q})}{\| \nabla \mathcal{S}_1(\mathbf{q}) \times \nabla \mathcal{S}_2(\mathbf{q}) \|} + \frac{0.01}{2!} (\beta \nabla \mathcal{S}_1(\mathbf{q}) + \gamma \nabla \mathcal{S}_2(\mathbf{q})) \tag{3.17}$$

Obtaining $\mathbf{q}_c$

Given the quadratic interpolant to the curve at $\mathbf{q}_p$ in (3.17), we refine its estimate to a point on the curve by generating a sequence of points $\mathbf{q}_1$, $\mathbf{q}_2$,

$\cdots \rightarrow \mathbf{q}_c$ with $\mathbf{q}_0 = \mathbf{q}_p$. The Newton method for the solution of $\mathcal{S}_j(\mathbf{r}(s)) = 0$, $j = 1, 2$ at $\mathbf{r}(s) = \mathbf{q}_k$ is

$$\nabla \mathcal{S}_j(\mathbf{q}_k) \cdot (\mathbf{q}_{k+1} - \mathbf{q}_k) = -\mathcal{S}_j(\mathbf{q}_k) \tag{3.18}$$

where k is the iteration count. Expressing $\mathbf{\Delta}_k \equiv \mathbf{q}_{k+1} - \mathbf{q}_k$ as a linear combination of $\mathbf{r}'(\mathbf{q}_k)$, $\nabla \mathcal{S}_1(\mathbf{q}_k)$ and $\nabla \mathcal{S}_2(\mathbf{q}_k)$ as

$$\mathbf{\Delta}_k = \varsigma_k \mathbf{r}'(\mathbf{q}_k) + \vartheta_k \nabla \mathcal{S}_1(\mathbf{q}_k) + \varphi_k \nabla \mathcal{S}_2(\mathbf{q}_k) \tag{3.19}$$

and substituting in (3.18) results in

$$\nabla \mathcal{S}_j(\mathbf{q}_k) \cdot (\varsigma_k \mathbf{r}'(\mathbf{q}_k) + \vartheta_k \nabla \mathcal{S}_1(\mathbf{q}_k) + \varphi_k \nabla \mathcal{S}_2(\mathbf{q}_k)) = -\mathcal{S}_j(\mathbf{q}_k)$$

The choice $\varsigma_k = 0$ leads to the following system of equations for ϑ_k and φ_k

$$\begin{bmatrix} \nabla \mathcal{S}_1(\mathbf{q}_k) \cdot \nabla \mathcal{S}_1(\mathbf{q}_k) & \nabla \mathcal{S}_1(\mathbf{q}_k) \cdot \nabla \mathcal{S}_2(\mathbf{q}_k) \\ \nabla \mathcal{S}_2(\mathbf{q}_k) \cdot \nabla \mathcal{S}_1(\mathbf{q}_k) & \nabla \mathcal{S}_2(\mathbf{q}_k) \cdot \nabla \mathcal{S}_2(\mathbf{q}_k) \end{bmatrix} \left\{ \begin{array}{c} \vartheta_k \\ \varphi_k \end{array} \right\} = - \left\{ \begin{array}{c} \mathcal{S}_1(\mathbf{q}_k) \\ \mathcal{S}_2(\mathbf{q}_k) \end{array} \right\} \tag{3.20}$$

Solution of (3.20) along with (3.19) is used to obtain $\mathbf{\Delta}_k$. The Newton method (3.18) is iterated until a convergence criterion is met as shown in Algorithm 4.

Algorithm 4 Predictor-Corrector scheme

```
function TRACE $(\mathbf{q}, \mathcal{S}_1, \mathcal{S}_2)$
    $\begin{Bmatrix} \beta \\ \gamma \end{Bmatrix} \leftarrow \begin{bmatrix} \nabla\mathcal{S}_1(\mathbf{q}) \cdot \nabla\mathcal{S}_1(\mathbf{q}) & \nabla\mathcal{S}_1(\mathbf{q}) \cdot \nabla\mathcal{S}_2(\mathbf{q}) \\ \nabla\mathcal{S}_2(\mathbf{q}) \cdot \nabla\mathcal{S}_1(\mathbf{q}) & \nabla\mathcal{S}_2(\mathbf{q}) \cdot \nabla\mathcal{S}_2(\mathbf{q}) \end{bmatrix}^{-1} \begin{Bmatrix} -\mathbf{r}'(\mathbf{q}) \cdot H_{\mathcal{S}_1}(\mathbf{q})\,\mathbf{r}'(\mathbf{q}) \\ -\mathbf{r}'(\mathbf{q}) \cdot H_{\mathcal{S}_2}(\mathbf{q})\,\mathbf{r}'(\mathbf{q}) \end{Bmatrix}$
    $\mathbf{q}_p \leftarrow \mathbf{q} + 0.1\frac{\nabla\mathcal{S}_1(\mathbf{q}) \times \nabla\mathcal{S}_2(\mathbf{q})}{\|\nabla\mathcal{S}_1(\mathbf{q}) \times \nabla\mathcal{S}_2(\mathbf{q})\|} + \frac{0.01}{2}\left(\beta\nabla\mathcal{S}_1(\mathbf{q}) + \gamma\nabla\mathcal{S}_2(\mathbf{q})\right)$    ▷ Second order approximant
    $k \leftarrow 0$
    $\Delta_0 \leftarrow \mathbf{q}_0 \leftarrow \mathbf{q}_p$    ▷ Initial guess to the Newton method is the second order approximant
    while $(\|\Delta_k\| > 10^{-6}\|\mathbf{q}_k\|)$ do    ▷ Newton Loop
        $\begin{Bmatrix} \vartheta_k(\mathbf{q}_k) \\ \varphi_k(\mathbf{q}_k) \end{Bmatrix} \leftarrow \begin{bmatrix} \nabla\mathcal{S}_1(\mathbf{q}_k) \cdot \nabla\mathcal{S}_1(\mathbf{q}_k) & \nabla\mathcal{S}_1(\mathbf{q}_k) \cdot \nabla\mathcal{S}_2(\mathbf{q}_k) \\ \nabla\mathcal{S}_2(\mathbf{q}_k) \cdot \nabla\mathcal{S}_1(\mathbf{q}_k) & \nabla\mathcal{S}_2(\mathbf{q}_k) \cdot \nabla\mathcal{S}_2(\mathbf{q}_k) \end{bmatrix}^{-1} \begin{Bmatrix} -\mathcal{S}_1(\mathbf{q}_k) \\ -\mathcal{S}_2(\mathbf{q}_k) \end{Bmatrix}$
        $\Delta_k \leftarrow \vartheta_k(\mathbf{q}_k)\nabla\mathcal{S}_1(\mathbf{q}_k) + \varphi_k(\mathbf{q}_k)\nabla\mathcal{S}_2(\mathbf{q}_k)$
        $\mathbf{q}_{k+1} \leftarrow \mathbf{q}_k + \Delta_k$
        $k \leftarrow k + 1$
    $\mathbf{q}_c \leftarrow \mathbf{q}_{k+1}$
    $\mathbf{q} \leftarrow \mathbf{q}_c$
```

3.5.3 Obtaining volume of intersection polyhedron using Delaunay triangulation

We use a library code TetGen (see Si [77]) for this purpose. The library code takes as input the coordinates of the set of points on the periphery of intersection polyhedron and decomposes the polyhedron into multiple 3-simplices or tetrahedra. This process is refered to as Delaunay triangulation. Let $\mathcal{D}_{\mathcal{E}}$ be the Delaunay triangulation of $\mathcal{E}$ consisting of tetrahedra T such that $\mathcal{E} = \bigcup\limits_{T \in \mathcal{D}_{\mathcal{E}}} T$. Denoting $\mathbf{v}_0$, $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ as position vectors of the vertices of T and $(\mathbf{v}_1 - \mathbf{v}_0)$, $(\mathbf{v}_2 - \mathbf{v}_0)$ and $(\mathbf{v}_3 - \mathbf{v}_0)$ as columns of the 3×3 matrix $\mathcal{X}^T$, we get

$$Meas(\mathcal{E}) = \sum_{T \in \mathcal{D}} Meas(T) = \sum_{T \in \mathcal{D}} \frac{1}{3!} det(\mathcal{X}^T)$$

3.6 Summary of steps involved in construction of mapping operators

The process of obtaining the mapping operators is summarized here

- Identify each pair of overlapping flow and poromechanics elements.
- For each such pair, obtain the equations of the faces of the flow and poromechanics element.
- Obtain the points on the periphery of the intersection polyhedron of the aforementioned pair using surface-surface intersections.
- With points on the periphery of the intersection polyhedron as input, obtain the volume of the intersection polyhedron using Delaunay triangulation.
- Use the obtained volume to construct the local mapping operators.

3.7 Verification: Mandel's problem solution for consolidation in homogeneous porous medium

Mandel [56] solved the Biot equations for the case of a rectangular porous specimen with incompressible pore fluid in plane strain conditions

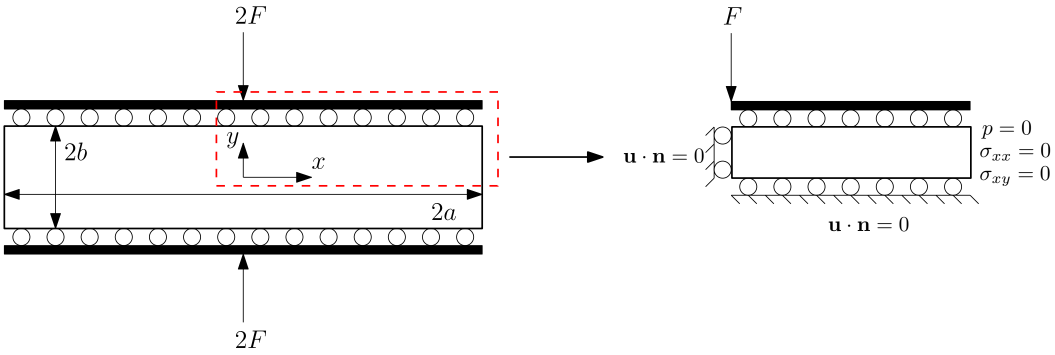

Figure 3.7: Circles indicate rollers and solid black boxes indicate rigid frictionless plates. The biaxial symmetry of the problem allows us to replicate the problem by only modeling a quarter of the domain as indicated by the red dotted line.

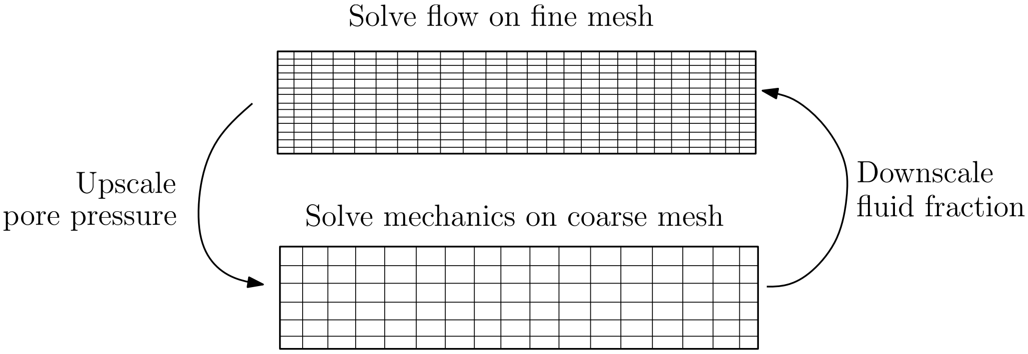

Figure 3.8: Solution methodology

sandwiched between frictionless plates and subjected to a point load with free drainage on the lateral sides. He arrived at a non-monotonic pore pressure solution at the center of a specimen i.e. an increase in pore pressure from $t = 0^+$ until a certain time at which the pressure reached a maximum followed by a monotonic decrease. Cryer [25] arrived at a similar non-monotonic pore pressure response at the center of water saturated sphere of soil subjected to uniform hydrostatic pressure with free drainage at the surface. As a result,

Parameter	Quantity	Value
a	x dimension	$100\,m$
b	y dimension	$10\,m$
E	Young's Modulus	$5.94 \times 10^9\,Pa$
ν	Poisson's ratio	0.2
ν_u	Undrained Poisson's ratio	0.3846
α	Biot parameter	0.8
k	Permeability	$100\,md$
B	Skempton coefficient	0.8333
c	Fluid compressibility	$3.03 \times 10^{-10}\,Pa^{-1}$
ϕ_0	Initial porosity	0.2
μ	Fluid viscosity	$1.0\,cp$
ρ_0	Reference fluid density	$62.4\,lbm/ft^3$
F	Point load intensity	$5.94 \times 10^8\,N/m$

Table 3.1: Parameters for Mandel's problem

$\mathscr{T}_h^f$	$\mathscr{T}_H^p$	$\\|\frac{1}{M}(p - p_H)\\|_{L^\infty(L^2)}$	Rate
15×15	6×6	0.459×10^{-1}	-
20×20	8×8	0.339×10^{-1}	1.053
25×25	10×10	0.265×10^{-1}	1.104
30×30	12×12	0.213×10^{-1}	1.198

Table 3.2: Order of convergence of upscaled pore pressure solution using the two grid fixed stress split iterative scheme for the Mandel's problem.

the phenomenon is commonly refered to as the 'Mandel-Cryer effect'. He observed that a monotonic pore pressure response is obtained only in the singular case when the shear modulus of the skeleton is zero. He also used the three-dimensional Terzaghi theory to arrive at the pore pressure response for the same configuration and observed that in that case, the water pressure drops steadily. He also showed that when the Poisson's ratio of the skeleton is 0.5, the Biot model predicts zero volume change whereas the Terzaghi theory predicts non-zero volume changes. The Mandel-Cryer effect was later experimentally verified by Gibson et al. [38] and Verruijt [89]. The reproducibility of this effect was the primary response why the Biot theory gained popularity in comparison to the Terzaghi theory. Later, Abousleiman et al. [1] obtained an analytical solution for the Mandel's problem but with the pore fluid compressibility factored in. This solution commonly serves as a benchmark for validation of coupled flow and poroelasticity codes. In our two grid implementation, the flow and mechanics domains, although identical, have different finite element discretizations, with mechanics being resolved on a coarser mesh. For the sake of clarity, we write the governing equations applicable to the Mandel's problem here as follows

$$\nabla \cdot \Big(\boldsymbol{\sigma}_0 + \frac{1}{2}\mathbb{D}(\nabla \mathbf{u} + \nabla^T \mathbf{u}) - \alpha(p - p_0)\mathbf{I} \Big) = \mathbf{0} \tag{3.21}$$

$$\frac{\partial}{\partial t}\Big(\frac{1}{M}p + \nabla \cdot (\alpha \mathbf{u}) \Big) + \nabla \cdot \Big(-\frac{k}{\mu}\nabla p \Big) = 0 \tag{3.22}$$

where (3.21) is the usual linear momentum balance for the solid phase with the small strain assumption in the absence of gravity and (3.22) is the equation of mass conservation for linearized slightly compressible single phase flow with gravity turned off. As shown in Figure 5.2, an infinitely long rectangular isotropic specimen is sandwiched between rigid, frictionless plates. The lateral sides are free from normal and shear stress and pore pressure. At $t = 0^+$, a force intensity of $2F\, N/m$ is applied to the rigid plates. The initial and boundary conditions are

$$\sigma_{xx}|_{t=0} = \sigma_{xy}|_{t=0} = \sigma_{yy}|_{t=0} = 0, p|_{t=0} = 0 \qquad \forall x, y$$

$$\sigma_{xx}|_{x=\pm a} = \sigma_{xy}|_{x=\pm a} = \sigma_{yx}|_{y=\pm b} = 0, \Big(\int_{-a}^{a} \sigma_{yy} dx \Big)_{y=\pm b} = -2F \qquad \forall t$$

$$p|_{x=\pm a} = 0, \big(\mathbf{u} \cdot \mathbf{n}\big)_{y=\pm b} = 0 \qquad \forall t$$

where $\mathbf{n}$ is unit outward normal to the boundary. Plane strain condition is applicable i.e. $\epsilon_{zz} = 0$. Given the biaxial symmetry of the problem, only a quarter of the domain needs to be modeled. Following the approach of Mikelić et al. [59], with $U_y^a(b)$ representing the analytical solution for the y displacement at $y = b$, the boundary conditions are recast as

$$\sigma_{xx}|_{x=a} = \sigma_{xy}|_{x=a} = \sigma_{yx}|_{y=b} = 0, \big(\mathbf{u} \cdot \mathbf{n}\big)_{y=b} = U_y^a(b) \qquad \forall t \tag{3.23}$$

$$\big(\mathbf{u} \cdot \mathbf{n}\big)_{x=0} = \big(\mathbf{u} \cdot \mathbf{n}\big)_{y=0} = 0,\ p|_{x=a} = 0 \qquad \forall t$$

$$\big(\mathbf{u} \cdot \mathbf{n}\big)_{x=0} = \big(\mathbf{u} \cdot \mathbf{n}\big)_{y=0} = \big(\mathbf{u} \cdot \mathbf{n}\big)_{y=b} = 0 \qquad \forall t$$

We solve the system (3.22)-(3.21) using the two grid scheme on rectilinear non-matching grids as shown in Figure 3.8 and show its convergence by measuring the upscaled pressure solution error. We employ the parameters given in Table 3.1 and keep the mesh ratio r fixed. Since the inverse of the Biot modulus $\frac{1}{M}$ is bounded below by a positive constant (as the initial porosity field ϕ_0 is strictly positive), optimality should be achieved when the pressure solution error is measured in the $L^\infty(L^2)$ norm (see Phillips and Wheeler [68]). The error norm is computed using the midpoint quadrature rule: $\|\frac{1}{M}(p - p_H)\|_{L^\infty(L^2)} \equiv \max_{0<\tau\leq T} \left(\sum_{E \in \mathcal{T}_H^p} |E| \left(\frac{p(\tau, m_e) - p_H(m_e)}{M} \right)^2 \right)^{\frac{1}{2}}$ where m_e is the center of mass of element E and T is the total time. To minimize the effects of the error produced by time discretization, a small time step of $1 \times 10^{-3}\, sec$ is chosen. As shown in

Table 5.1, we observe first order convergence for the pore pressure solution with the augmented scheme for $r = 2.5$. A fractional value of r ensures the cardinality $|\{\mathcal{E}\}|$ of the set $\{\mathcal{E}\}$ of intersection polyhedra is non-zero i.e. the number of instances of intersecting flow and mechanics elements is not zero. It is important to note that there is no restriction posed on the value of r and we choose a value of 2.5 only for the sake of convenience. We then compare the upscaled pore pressure solution at the cell-center closest to the origin of the quarter domain with the analytical solution for all the above combinations of $\mathscr{T}_h^f$, $\mathscr{T}_H^p$. The reason for choosing the cell-center closest to the origin of the quarter domain is that the classical non-monotonic pore pressure response, which we intend to replicate in our numerical model, is expected only near the central region of the specimen (see Abousleiman et al. [1]). We also compare the computed x-displacement at the free end $x = a$ with the analytical solution. The total simulation time is $50000\,sec$ with a time step of $10\,sec$. According to the analytical solution, at the instant of loading, a uniform pressure rise of $\Delta p(x, y, 0^+) = \frac{FB(1+\nu_u)}{3a}$ should be observed. Also, after the initial outward movement of $u_x(a, y, 0^+) = \frac{F\nu_u}{2G}$, the side boundaries will contract toward the center and its final state should be $u_x(a, y, \infty) = \frac{F\nu}{2G}$. The pore pressure and displacement solutions are non-dimensionalized by multiplying with $(\frac{a}{F})$ and $(\frac{2G}{F})$ respectively. As shown in Figures 3.9 and 3.10, we observe an excellent match with the expected results for all the above combinations.

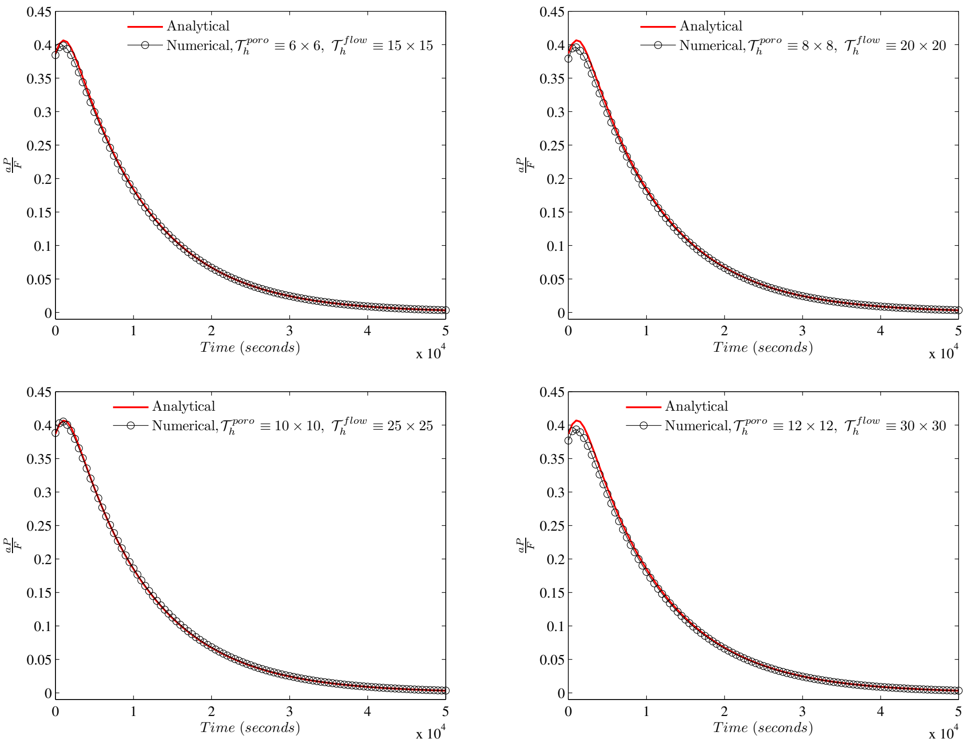

Figure 3.9: Non-monotonic pore pressure response at the cell-center closest to the origin for the Mandel's problem with nonmatching grids. $\lim_{t \to 0^+} \left(\frac{aP(x_c, y_c, t)}{F} \right) = \frac{B(1+\nu_u)}{3} = 0.3846$ where x_c, y_c are coordinates of cell-center closest to the origin.

3.8 Large-scale problem

As shown in Figure 3.11, the large-scale problem under consideration consists of a reservoir of dimensions $80\,ft \times 9400\,ft \times 8800\,ft$ at a depth of approximately $10000\,ft$ with a dip-angle ranging from 1^o to 3^o surrounded by a geomechanical domain of dimensions $13000\,ft \times 9400\,ft \times 8800\,ft$ with an underburden of approximately $3000\,ft$. Sideburdens are neglected for the sake

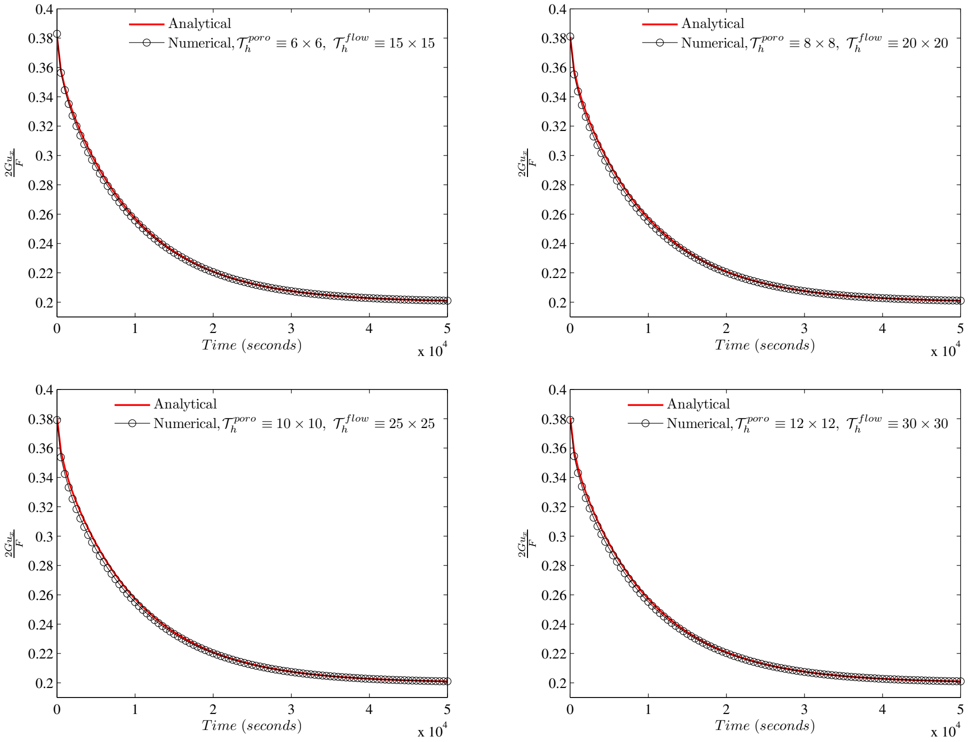

Figure 3.10: Displacement response at the free end for the Mandel's problem with nonmatching grids. $\lim_{t\to 0^+}\left(\frac{2Gu_x(a,y,t)}{F}\right) = \nu_u = 0.3846$ and $\lim_{t\to\infty}\left(\frac{2Gu_x(a,y,t)}{F}\right) = \nu = 0.2$.

of convenience. The flow mesh $\mathscr{T}_h^f \equiv 20\times188\times176$ consists of 661760 distorted hexahedral elements of size roughly $4\,ft \times 50\,ft \times 50\,ft$. As explained in Table 3.3, the graded geomechanics mesh $\mathscr{T}_H^p \equiv 82 \times 47 \times 47$ consists of 169576 brick elements. At the outset, it is obvious that in the absence of a multiscale framework, the geomechanics mesh with similar resolution for overburden and underburden would need to conform with the flow grid and would be roughly $\mathscr{T}_H^p \equiv (20+33+7)\times 188 \times 176$ consisting of 1654400 elements, which

Nature	Depthwise extent	Mesh	Element size
Overburden	$0\,ft \rightarrow 9900\,ft$	$33 \times 47 \times 44$	$300\,ft \times 200\,ft \times 200\,ft$
Near-reservoir	$9900\,ft \rightarrow 10300\,ft$	$40 \times 47 \times 44$	$10\,ft \times 200\,ft \times 200\,ft$
Underburden	$10300\,ft \rightarrow 13000\,ft$	$9 \times 47 \times 44$	$300\,ft \times 200\,ft \times 200\,ft$

Table 3.3: Details of the graded geomechanics mesh

is an order of magnitude higher than the number of geomechanics elements currently employed. The flow domain is subject to no flux on all its boundaries whereas the geomechanics domain has a traction-free top surface and zero normal displacements on the remaining boundaries. The fluid compressibility is $1.45 \times 10^{-8}\,Pa^{-1}$, the fluid viscosity is $1\,cp$ and the initial reservoir porosity is taken to be 0.2. The Biot parameter is taken to be 0.8 for the near-reservoir geomechanics elements and zero for the elements discretizing the overburden and underburden. The Young's modulus is taken to be $5.94 \times 10^{9}\,Pa$ and the Poisson's ratio is 0.2. The initial reservoir hydrostatic pressure is taken to be $4650\,psi$, the reservoir permeability is taken to be $100\,md$ and the rock density is taken to be $2.65\,g/cm^3$. An injection well of BHP $7000\,psi$ is located at $y = 2000\,ft, z = 2000\,ft$ whereas a production well of BHP $2000\,psi$ is located at $y = 6000\,ft, z = 2000\,ft$.

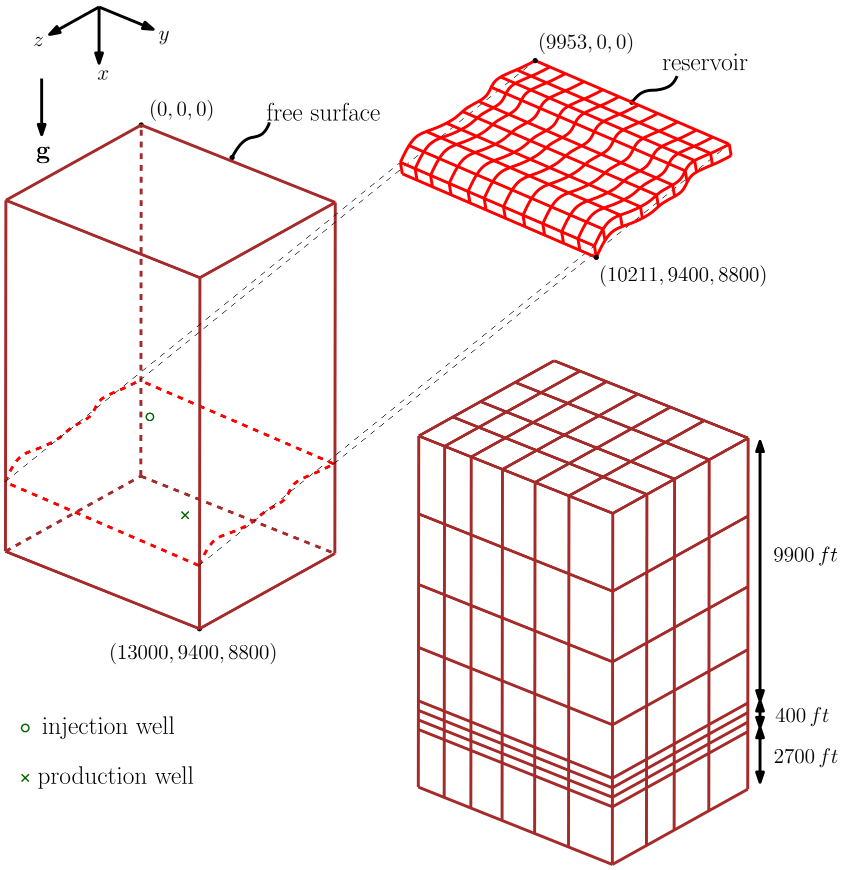

Figure 3.11: Flow domain of $80\,ft \times 9400\,ft \times 8800\,ft$ surrounded by a geomechanics domain of $13000\,ft \times 9400\,ft \times 8800\,ft$. Injection well is located at $y = 2000\,ft, z = 2000\,ft$ and production well is located at $y = 6000\,ft, z = 6000\,ft$.

Figures 3.12, 3.13 and 3.14 are the results of the simulation on 128 processes with a time step of $1\,day$ at the end of $200\,days$. The total run time was $15\,minutes$. The injection well generates an increase in pore pressure from the in situ hydrostatic pressure of $4650\,psi$ whereas the production well generates a decrease in pore pressure with radial flow patterns as shown in

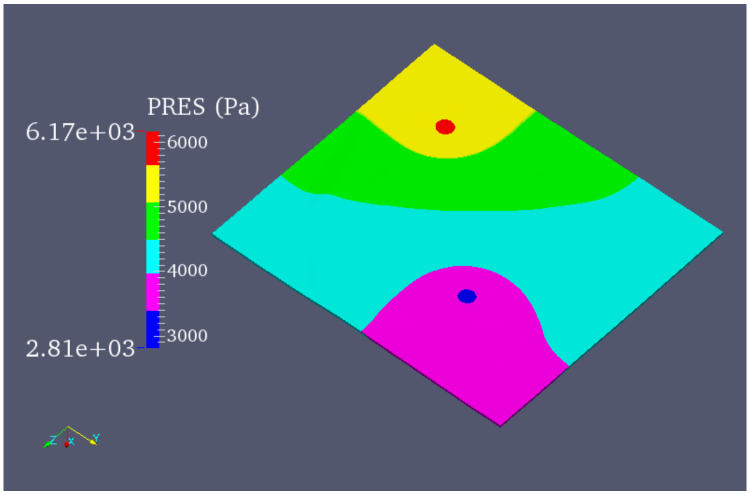

Figure 3.12: Reservoir pore pressure distribution at the end of 200 $days$.

Figure 3.12. The concomitant geomechanical response is shown in Figure 3.13. As the x axis points vertically downwards, positive x displacements indicate movement in the direction of gravity whereas negative x displacements indicate movement against the direction of gravity. The distribution of the vertical displacement on the free surface approximately 10000 ft above the reservoir is shown in Figure 3.14. This distribution follows the radial flow patterns generated from the wells, while keeping in mind that the effect of fluid injection in the reservoir buried deep beneath the free surface is to induce an uplift whereas the effect of fluid withdrawal is to induce subsidence. As a result, portions of the geomechanical domain closer to the production well laterally experience larger movements in the direction of gravity resulting in larger positive values for x displacement.

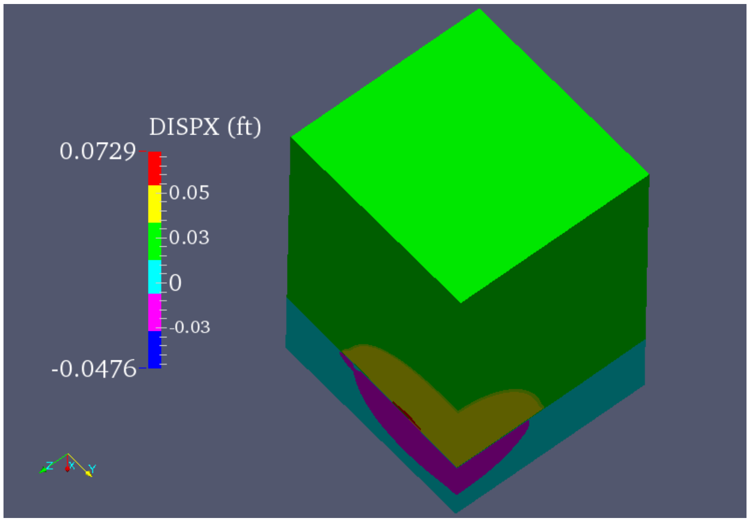

Figure 3.13: Vertical displacement distribution in the geomechanical domain at the end of 200 *days*.

3.9 Summary

Since flow and poromechanics are solved on different finite element grids, we design grid-to-grid projection operators using concepts in computational geometry and parallel computing. We first carve out a code framework in which flow is solved in parallel on a MPI subcommunicator and poromechanics is solved on the default MPI communicator. We then look at a pair of geometrically non-disjoint flow finite element and poromechanics finite element as two three-dimensional objects intersecting one another. We then invoke singular value decompositions to obtain the equations of the surfaces of the flow and poromechanics elements. We then subject those equations to a predictor-corrector algorithm in order to perform surface-surface intersections. These intersections are used to obtain points on the periphery of the

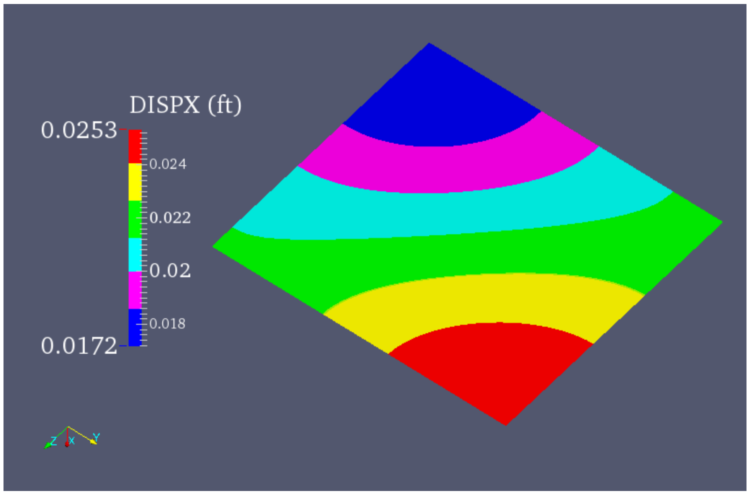

Figure 3.14: Vertical displacement distribution at the free surface at the end of 200 $days$.

intersection polyhedron of that pair. We then employ a Delaunay triangulation routine that takes these points as input and provides the tetrahedralization of the intersection polyhedron. We add up the volumes of the tetrahedra to obtain the volume of the intersection polyhedron. This volume is then used to implement local grid-to-grid projection operators. The global grid-to-grid projection operators would be the assembly of the local operators. This entire code machinery is impemented as an addendum to the existing staggered solution algorithmic code framework that solves flow and poromechanics on the same finite element grid. We then establish the numerical convergence of the solution algorithm for the benchmark problem, and demonstrate its capability in handling a large-scale field problem.

Chapter 4

Convergence analysis of the two-grid algorithm and the link with computational homogenization for heterogeneous poroelastic media

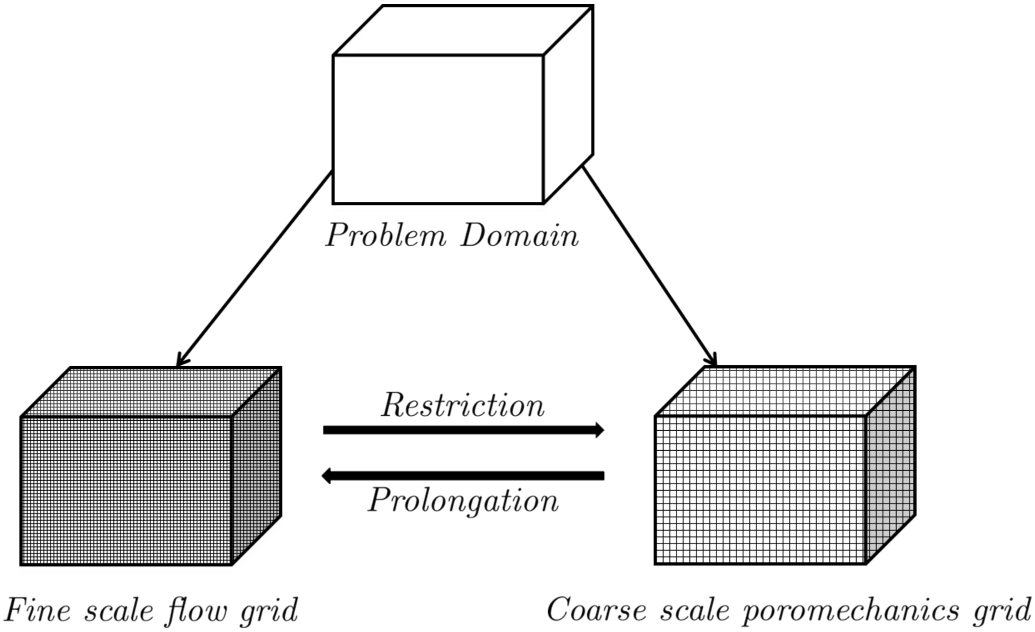

Figure 4.1: One coupling iteration of two-grid scheme. After the flow solve, the updated variables are projected onto the coarse scale mechanics grid. After the mechanics solve, the updated variables are projected onto the fine scale flow grid. In order to be consistent with the terminology used in multigrid methods, we refer to projection onto coarse grid as ‘restriction’ and projection onto fine grid as ‘prolongation’.

In chapters 2 and 3, we developed a two grid fixed stress split scheme in which the flow equations are solved on a fine grid while freezing the total mean stress of the porous solid and the poromechanics equations are solved on a coarse grid, with the grids being non-nested. Operators are constructed in the pre-processing stage which map the flow solution to the coarse grid and the poromechanical solution to the fine grid in every coupling iteration at every time step. We finally used the classical Mandel's problem analytical solution to show that the scheme is numerically convergent. The objective of the work in this chapter is to establish convergence of the two grid scheme from a theoretical standpoint. It is important to note that while the work presented in chapters 2 and 3 applies to non-nested grids for slightly compressible single phase flow coupled with linear poromechanics with the flow domain being a subset of the poromechanical domain, the theoretical convergence analysis presented in this chapter applies to the degenerate case of nested grids for the *linearized flow model* coupled with linear poromechanics with the flow and poromechanical domains being identical, as shown in Figure 4.1. The convergence analysis is motivated by the previous work of Mikelić and Wheeler [58] and Almani et al. [3]. Mikelić and Wheeler [58] proved that the standard fixed stress split scheme for the linearized flow model coupled with linear poromechanics in which the subproblems are resolved on the same time and length scales is a contraction map with respect to appropriately chosen metrics. Almani et al. [3] extended those results to establish convergence of the multirate fixed stress split scheme in which the subproblems are resolved on the same length scale but different time scales with the poromechanics subproblem being resolved on a larger time scale.

4.1 Flow model

The fluid mass conservation equation (4.1) in the presence of deformable porous medium with the Darcy law (4.2) and linear pressure dependence of density (4.3) with boundary conditions (4.4) and initial conditions (4.5) is

$$\frac{\partial \zeta}{\partial t} + \nabla \cdot \mathbf{z} = q \tag{4.1}$$

$$\mathbf{z} = -\frac{\mathbf{K}}{\mu}(\nabla p - \rho \mathbf{g}) = -\boldsymbol{\kappa}(\nabla p - \rho \mathbf{g}) \tag{4.2}$$

$$\rho = \rho_0(1 + c\,(p - p_0)) \tag{4.3}$$

$$p = g \text{ on } \Gamma_D^f \times (0,T], \; \mathbf{z}\cdot\mathbf{n} = 0 \text{ on } \Gamma_N^f \times (0,T] \tag{4.4}$$

$$p(\mathbf{x},0) = p_0(\mathbf{x}), \; \rho(\mathbf{x},0) = \rho_0(\mathbf{x}), \; \phi(\mathbf{x},0) = \phi_0(\mathbf{x}) \qquad (\forall \mathbf{x} \in \Omega) \tag{4.5}$$

where $p : \Omega \times (0,T] \to \mathbb{R}$ is the fluid pressure, $\mathbf{z} : \Omega \times (0,T] \to \mathbb{R}^3$ is the fluid flux, $\bar{\epsilon}$ is the volumetric strain, $\zeta \equiv \frac{1}{M}p + \alpha\bar{\epsilon}$ is the fluid content, Γ_D^f is the Dirichlet boundary, $\mathbf{n}$ is the unit outward normal on the Neumann boundary Γ_N^f, q is the source or sink term[1], $\mathbf{K}$ is the uniformly symmetric positive definite absolute permeability tensor, μ is the fluid viscosity, ρ_0 is a reference density, $\boldsymbol{\kappa} = \frac{\mathbf{K}}{\mu}$ is a measure of the hydraulic conductivity of the pore fluid, c is the fluid compressibility, $T > 0$ is the time interval, α is the Biot constant and

---

[1] It is important to note that this term has the units of $[M^0L^0T^{-1}]$ whereas the source or sink term in Equation (2.1) in Chapter 2 for the slightly compressible single phase flow model has the units of $[M^1L^{-3}T^{-1}]$

$M \equiv \left(\phi_0 c + \frac{(\alpha-\phi_0)(1-\alpha)}{K_b}\right)^{-1}$ is the Biot modulus with K_b being the drained bulk modulus.

4.1.1 Arriving at the mass conservation equation for the linearized flow model

We first comment that the difference between the slightly compressible single phase flow model and the linearized flow model lies in the expression for the density variation as a function of pore pressure. The former models the density variation as

$$\rho = \rho_0 e^{c(p-p_0)}$$

which when expanded as a Taylor series, is written as

$$\rho = \rho_0(1 + c(p-p_0) + \frac{c^2(p-p_0)^2}{2!} + \cdots)$$

which, after noting that the fluid compressibility is of the order of 10^{-6} can be approximated as

$$\rho \approx \rho_0(1 + c(p-p_0)) \tag{4.6}$$

after truncating the expansion to the first order term. The above expression, identical to (4.3) is the function for variation of density as a function of pore pressure for the *linearized* flow model. We also note that while the slightly compressible single phase flow might require a nonlinear solve with Newton iterations within every coupling iteration, the linearized flow model requires only one solve within every coupling iteration. Furthermore, the Lagrangian porosity variation, instead of having two different expressions for the flow solve and the poromechanics solve as we saw in modules 2.3.7 and 2.3.8 respectively, has only one expression given by

$$\phi^* = \phi_0 + \alpha\bar{\epsilon} + \frac{1}{N}(p - p_0) \tag{4.7}$$

with

$$\frac{1}{N} \equiv \frac{(\alpha - \phi_0)(1 - \alpha)}{K_b}$$

in lieu of Equation (2.47). We start with the mass conservation for slightly compressible single flow in porous medium as follows

$$\frac{\partial(\phi^* \rho)}{\partial t} - \nabla \cdot (\rho \boldsymbol{\kappa}(\nabla p - \rho \mathbf{g})) = q \tag{4.8}$$

It is again important to note that the term q on the RHS has the units of $[M^1 L^{-3} T^{-1}]$. We substitute (4.6) and (4.7) in (4.8) to get

$$\begin{aligned}\frac{\partial((\phi_0 + \alpha\bar{\epsilon} + \frac{1}{N}(p - p_0))\rho_0(1 + c(p - p_0)))}{\partial t}\\ - \nabla \cdot (\rho_0(1 + c(p - p_0))\boldsymbol{\kappa}(\nabla p - \rho \mathbf{g})) = q\end{aligned}$$

which, after dividing by ρ_0, is obtained as

$$\begin{aligned}\frac{\partial((\phi_0 + \alpha\bar{\epsilon} + \frac{1}{N}(p - p_0))(1 + c(p - p_0)))}{\partial t}\\ - \nabla \cdot (\boldsymbol{\kappa}(1 + c(p - p_0))(\nabla p - \rho \mathbf{g})) = \frac{q}{\rho_0}\end{aligned}$$

which is written as

$$\frac{\partial(\phi_0 + \alpha\bar{\epsilon} + \frac{1}{N}(p - p_0))}{\partial t} + \frac{\partial((\phi_0 + \alpha\bar{\epsilon})c(p - p_0))}{\partial t} + \frac{\partial(\frac{1}{N}(p - p_0)c(p - p_0))}{\partial t}$$

$$- \nabla \cdot (\boldsymbol{\kappa}(1 + c(p - p_0))(\nabla p - \rho \mathbf{g})) = \frac{q}{\rho_0}$$

which is further written as

$$\frac{\partial(\phi_0 + \alpha\bar{\epsilon} + \frac{1}{N}(p - p_0))}{\partial t} + (\phi_0 + \alpha\bar{\epsilon})c\frac{\partial(p - p_0)}{\partial t} + c(p - p_0)\frac{\partial(\phi_0 + \alpha\bar{\epsilon})}{\partial t}$$
$$+ \frac{\partial(\frac{1}{N}(p - p_0)c(p - p_0))}{\partial t} - \nabla \cdot (\boldsymbol{\kappa}(1 + c(p - p_0))(\nabla p - \rho \mathbf{g})) = \frac{q}{\rho_0}$$

which, in lieu of the knowledge that $\bar{\epsilon}$ is of the order of 10^{-3} and $\phi_0 \approx 0.2$ resulting in $\alpha\bar{\epsilon} \ll \phi_0$, and furthermore $\frac{\partial \phi_0}{\partial t} = 0$ and $\frac{\partial p_0}{\partial t} = 0$, is written as

$$\frac{\partial(\alpha\bar{\epsilon} + \frac{1}{N}p)}{\partial t} + \frac{\partial(\phi_0 c p)}{\partial t} + \frac{2c}{N}(p - p_0)\frac{\partial p}{\partial t} + \alpha c(p - p_0)\frac{\partial \bar{\epsilon}}{\partial t}$$
$$- \nabla \cdot (\boldsymbol{\kappa}(1 + c(p - p_0))(\nabla p - \rho \mathbf{g})) = \frac{q}{\rho_0}$$

We neglect the nonlinear terms $\frac{2c}{N}(p - p_0)\frac{\partial p}{\partial t}$ and $\alpha c(p - p_0)\frac{\partial \bar{\epsilon}}{\partial t}$ to obtain

$$\frac{\partial(\alpha\bar{\epsilon} + \frac{1}{N}p + \phi_0 c p)}{\partial t} - \nabla \cdot (\boldsymbol{\kappa}(1 + c(p - p_0))(\nabla p - \rho \mathbf{g})) = \frac{q}{\rho_0}$$

We expand the second term on the LHS to obtain

$$\frac{\partial(\alpha\bar{\epsilon} + \frac{1}{N}p + \phi_0 c p)}{\partial t} - \nabla \cdot (\boldsymbol{\kappa}(\nabla p - \rho \mathbf{g})) - \nabla \cdot (\boldsymbol{\kappa} c(p - p_0)(\nabla p - \rho \mathbf{g})) = \frac{q}{\rho_0}$$

We neglect the nonlinear term $\nabla \cdot (\boldsymbol{\kappa} c(p - p_0)(\nabla p - \rho \mathbf{g}))$ to obtain

$$\frac{\partial \zeta}{\partial t} - \nabla \cdot (\boldsymbol{\kappa}(\nabla p - \rho \mathbf{g})) = \frac{q}{\rho_0}$$

where

$$\zeta = \alpha\bar{\epsilon} + \frac{1}{N}p + \phi_0 c p \equiv \alpha\bar{\epsilon} + \frac{1}{M}p$$

is refered to as the fluid content and

$$\frac{1}{M} = \frac{1}{N} + \phi_0 c \equiv \phi_0 c + \frac{(\alpha - \phi_0)(1 - \alpha)}{K_b}$$

4.2 Poromechanics model

The equations of the poromechanics model are already given in Module 2.1.2.

4.3 Statement of contraction of the fully discrete two-grid fixed stress split scheme

The basic idea of the two-grid fixed stress split iterative strategy is to solve the flow system (4.1)-(4.5) on a fine grid for the pressures at the current coupling iteration based on the value of the mean stress from the previous coupling iteration. These pressures are then fed to the poromechanics system (2.6)-(2.11) which is solved for displacements on a coarse grid thereby updating the stress state. This updated stress state is then fed back to the flow system for the next coupling iteration. Since this strategy condemns the porous solid to follow a certain stress path during the flow solve, the convergence of the solution algorithm is not automatically guaranteed. The objective of our analysis is to arrive at a contraction map for the fully discrete two-grid fixed stress split iterative scheme while taking into account the heterogeneities in the underlying porous medium. Let $\mathscr{T}_h^f$ represent the fine scale flow grid consisting of brick elements E^f and $\mathscr{T}_H^p$ be the coarse scale poromechanical grid consisting of brick elements E^p such that

$$r = \max_{E^p \in \mathscr{T}_H^p} diam(E^p) / \max_{E^f \in \mathscr{T}_h^f} diam(E^f) = H/h \geq 1$$

Since the grids ar nested, each coarse scale poromechanical element $E^p \in \mathscr{T}_H^p$ can be viewed as a union of flow elements belonging to the set $\mathfrak{I}^{E^p}$ as follows

$$E^p = \bigcup_{E^f \in \mathfrak{I}^{E^p}} E^f \qquad \text{where} \qquad \mathfrak{I}^{E^p} \equiv \left\{ E^f : E^f \subset E^p \; \forall E^f \in \mathfrak{T}_h^f \right\}$$

We use the notations $(\cdot)^{n+1}$ for any quantity $(\cdot)$ evaluated at time level $n+1$, $(\cdot)^{m,n+1}$ for any quantity $(\cdot)$ evaluated at the m^{th} coupling iteration at time level $n+1$, $\delta_f^{(m)}(\cdot)$ for the change in the quantity $(\cdot)$ during the flow solve over the $(m+1)^{th}$ coupling iteration at any time level and $\delta^{(m)}(\cdot)$ for the change in the quantity $(\cdot)$ over the $(m+1)^{th}$ coupling iteration at any time level. These notations shall be carried forward to Chapter 5 as well. Further, to take into account the underlying heterogeneities in the porous medium, we introduce the notations $(\cdot)_{E^f}$ for the value of any material parameter $(\cdot)$ at flow element E^f and $(\cdot)_{E^p}$ for the value of any material parameter $(\cdot)$ evaluated at poromechanics element E^p.

4.3.1 Discrete variational statements for the flow subproblem

Before arriving at the discrete variational statement of the flow model, we impose the fixed mean stress constraint on the strong form of the mass conservation equation (4.1). Invoking the relation for the total mean stress $\sigma_v = K_b\bar{\epsilon} - \alpha p$, we get

$$\frac{\partial}{\partial t}\left(\frac{1}{M}p + \alpha\left(\frac{\sigma_v + \alpha p}{K_b}\right)\right) + \nabla \cdot \mathbf{z} = q$$

$$\overbrace{\left(\frac{1}{M} + \frac{\alpha^2}{K_b}\right)}^{\varphi} \frac{\partial p}{\partial t} + \nabla \cdot \mathbf{z} = q - \frac{\alpha}{K_b}\frac{\partial \sigma_v}{\partial t} \tag{4.9}$$

Using backward Euler in time, the discrete in time form of (5.4) for the m^{th} coupling iteration in the $(n+1)^{th}$ time step is written as

$$\varphi \frac{1}{\Delta t}(p^{m,n+1} - p^n) + \nabla \cdot \mathbf{z}^{n+1} = q^{n+1} - \frac{\alpha}{K_b} \frac{1}{\Delta t}(\sigma_v^{m,n+1} - \sigma_v^n)$$

where Δt is the time step and the source term as well as the terms evaluated at the previous time level n do not depend on the coupling iteration count as they are known quantities. *The fixed mean stress constraint implies that* $\sigma_v^{m,n+1}$ *gets replaced by* $\sigma_v^{m-1,n+1}$ *i.e. the computation of* $p^{m,n+1}$ *and* $\mathbf{z}^{m,n+1}$ *is based on the value of the mean stress updated after the poromechanics solve from the previous coupling iteration* $m-1$ *at the current time level* $n+1$. The modified equation is written as

$$\varphi(p^{m,n+1} - p^n) + \Delta t \nabla \cdot \mathbf{z}^{m,n+1} = \Delta t q^{n+1} - \frac{\alpha}{K_b}(\sigma_v^{m-1,n+1} - \sigma_v^n) \tag{4.10}$$

As a result, the discrete variational statement of (4.1) in the presence of medium heterogeneities is

$$\begin{aligned} &\sum_{E^f \in \mathscr{T}_h^f} \varphi_{E^f}(p_h^{m,n+1} - p_h^n, \theta_h)_{E^f} + \sum_{E^f \in \mathscr{T}_h^f} \Delta t(\nabla \cdot \mathbf{z}_h^{m,n+1}, \theta_h)_{E^f} \\ &= \sum_{E^f \in \mathscr{T}_h^f} \Delta t(q^{n+1}, \theta_h)_{E^f} - \sum_{E^f \in \mathscr{T}_h^f} \frac{\alpha_{E^f}}{K_{b_{E^f}}}(\sigma_v^{m-1,n+1} - \sigma_v^n, \theta_h)_{E^f} \end{aligned} \tag{4.11}$$

Replacing m by $m+1$ in (4.11) and subtracting the two equations, we get

$$\begin{aligned} &\sum_{E^f \in \mathscr{T}_h^f} \varphi_{E^f}(\delta^{(m)} p_h, \theta_h)_{E^f} + \sum_{E^f \in \mathscr{T}_h^f} \Delta t(\nabla \cdot \delta^{(m)} \mathbf{z}_h, \theta_h)_{E^f} \\ &= - \sum_{E^f \in \mathscr{T}_h^f} \frac{\alpha_{E^f}}{K_{b_{E^f}}}(\delta^{(m-1)} \sigma_v, \theta_h)_{E^f} \end{aligned}$$

The weak form of the Darcy law (4.2) for the m^{th} coupling iteration in the $(n+1)^{th}$ time step is

$$(\boldsymbol{\kappa}^{-1}\mathbf{z}^{m,n+1}, \mathbf{v})_\Omega = -(\nabla p^{m,n+1}, \mathbf{v})_\Omega + (\rho_0 \mathbf{g}, \mathbf{v})_\Omega \qquad \forall\ \mathbf{v} \in \mathbf{V}(\Omega) \tag{4.12}$$

where $\mathbf{V}(\Omega)$ is given by

$$\mathbf{V}(\Omega) \equiv \mathbf{H}(div, \Omega) \cap \left\{\mathbf{v} : \mathbf{v} \cdot \mathbf{n} = 0 \text{ on } \Gamma_N^f\right\}$$

and $\mathbf{H}(div, \Omega)$ is given by

$$\mathbf{H}(div, \Omega) \equiv \left\{\mathbf{v} : \mathbf{v} \in (L^2(\Omega))^3, \nabla \cdot \mathbf{v} \in L^2(\Omega)\right\}$$

We use the divergence theorem to evaluate the first term on RHS of (5.5) as follows

$$\begin{aligned}(\nabla p^{m,n+1}, \mathbf{v})_\Omega &= (\nabla, p^{m,n+1}\mathbf{v})_\Omega - (p^{m,n+1}, \nabla \cdot \mathbf{v})_\Omega \\ &= (p^{m,n+1}, \mathbf{v} \cdot \mathbf{n})_{\partial\Omega} - (p^{m,n+1}, \nabla \cdot \mathbf{v})_\Omega = (g, \mathbf{v} \cdot \mathbf{n})_{\Gamma_D^f} - (p^{m,n+1}, \nabla \cdot \mathbf{v})_\Omega\end{aligned} \tag{4.13}$$

where we invoke $\mathbf{v} \cdot \mathbf{n} = 0$ on Γ_N^f. In lieu of (5.5) and (5.6), we get

$$(\boldsymbol{\kappa}^{-1}\mathbf{z}^{m,n+1}, \mathbf{v})_\Omega = -(g, \mathbf{v} \cdot \mathbf{n})_{\Gamma_D^f} + (p^{m,n+1}, \nabla \cdot \mathbf{v})_\Omega + (\rho_0 \mathbf{g}, \mathbf{v})_\Omega$$

As a result, the discrete variational statement of (4.2) in the presence of medium heterogeneities is

$$\begin{aligned}&\sum_{E^f \in \mathscr{T}_h^f} (\boldsymbol{\kappa}_{E^f}^{-1}\mathbf{z}_h^{m,n+1}, \mathbf{v}_h)_{E^f} - \sum_{E^f \in \mathscr{T}_h^f} (p_h^{m,n+1}, \nabla \cdot \mathbf{v}_h)_{E^f} \\ &= \sum_{E^f \in \mathscr{T}_h^f} (\rho_0 \mathbf{g}, \mathbf{v}_h)_{E^f} - \sum_{E^f \in \mathscr{T}_h^f} (g, \mathbf{v}_h \cdot \mathbf{n})_{\partial E^f \cap \Gamma_D^f}\end{aligned} \tag{4.14}$$

Replacing m by $m+1$ in (4.14) and subtracting the two equations, we get

$$\sum_{E^f \in \mathscr{T}_h^f} (\boldsymbol{\kappa}_{E^f}^{-1} \delta^{(m)} \mathbf{z}_h, \mathbf{v}_h)_{E^f} = \sum_{E^f \in \mathscr{T}_h^f} (\delta^{(m)} p_h, \nabla \cdot \mathbf{v}_h)_{E^f}$$

4.3.1.1 A note on the inconsistency for the flow solve during the first fixed stress split iteration

For the sake of clarity, we rewrite (4.10) here as follows

$$\varphi(p^{m,n+1} - p^n) + \Delta t \nabla \cdot \mathbf{z}^{m,n+1} = \Delta t q^{n+1} - \frac{\alpha}{K_b}(\sigma_v^{m-1,n+1} - \sigma_v^n)$$

and successively subsititute $m = 1, 2, 3, \cdots$ where m is the coupling iteration number to obtain

$$\left.\begin{aligned} &\varphi(p^{1,n+1} - p^n) + \Delta t \nabla \cdot \mathbf{z}^{1,n+1} = \Delta t q^{n+1} - \frac{\alpha}{K_b}(\sigma_v^{0,n+1} - \sigma_v^n) \\ &\varphi(p^{2,n+1} - p^n) + \Delta t \nabla \cdot \mathbf{z}^{2,n+1} = \Delta t q^{n+1} - \frac{\alpha}{K_b}(\sigma_v^{1,n+1} - \sigma_v^n) \\ &\varphi(p^{3,n+1} - p^n) + \Delta t \nabla \cdot \mathbf{z}^{3,n+1} = \Delta t q^{n+1} - \frac{\alpha}{K_b}(\sigma_v^{2,n+1} - \sigma_v^n) \\ &\cdots \end{aligned}\right\} \tag{4.15}$$

It is obvious that the mean stress $\sigma_v^{0,n+1}$ at the start of the current time step before any coupling iterations occur is the same as the mean stress σ_v^n at the end of the previous time step implying that

$$\sigma_v^{0,n+1} = \sigma_v^n$$

As a result, we write (4.15) as

$$\left.\begin{aligned} &\varphi(p^{1,n+1} - p^n) + \Delta t \nabla \cdot \mathbf{z}^{1,n+1} = \Delta t q^{n+1} \\ &\varphi(p^{2,n+1} - p^n) + \Delta t \nabla \cdot \mathbf{z}^{2,n+1} = \Delta t q^{n+1} - \frac{\alpha}{K_b}(\sigma_v^{1,n+1} - \sigma_v^n) \\ &\varphi(p^{3,n+1} - p^n) + \Delta t \nabla \cdot \mathbf{z}^{3,n+1} = \Delta t q^{n+1} - \frac{\alpha}{K_b}(\sigma_v^{2,n+1} - \sigma_v^n) \\ &\cdots \end{aligned}\right\} \tag{4.16}$$

Equation (4.16) clearly shows that the computation of the pressure $p^{1,n+1}$ during the flow solve in the first fixed stress interation does not factor in the mean stress term whereas computation of pressure during the flow solve in the subsequent fixed stress iterations does. This gives pointers to possible improvement in the fixed stress split algorithmic logic and provides scope for future work.

4.3.2 Discrete variational statement for the poromechanics subproblem

The weak form of the linear momentum balance (2.6) is given by

$$(\nabla \cdot \boldsymbol{\sigma}, \mathbf{q})_{\Omega} + (\mathbf{f} \cdot \mathbf{q})_{\Omega} = 0 \qquad (\forall\ \mathbf{q} \in \mathbf{U}(\Omega)) \tag{4.17}$$

where $\mathbf{U}(\Omega)$ is given by

$$\mathbf{U}(\Omega) \equiv \left\{\mathbf{q} = (u, v, w) : u, v, w \in H^1(\Omega), \mathbf{q} = \mathbf{0} \text{ on } \Gamma_D^p\right\}$$

where $H^m(\Omega)$ is defined, in general, for any integer $m \geq 0$ as

$$H^m(\Omega) \equiv \left\{w : D^{\alpha} w \in L^2(\Omega)\ \forall |\alpha| \leq m\right\},$$

where the derivatives are taken in the sense of distributions and given by

$$D^{\alpha} w = \frac{\partial^{|\alpha|} w}{\partial x_1^{\alpha_1} .. \partial x_n^{\alpha_n}},\ |\alpha| = \alpha_1 + \cdots + \alpha_n,$$

We know from tensor calculus that

$$(\nabla \cdot \boldsymbol{\sigma}, \mathbf{q}) \equiv (\nabla, \boldsymbol{\sigma}\mathbf{q}) - (\boldsymbol{\sigma}, \nabla \mathbf{q}) \tag{4.18}$$

Further, using the divergence theorem and the symmetry of $\boldsymbol{\sigma}$, we arrive at

$$(\nabla, \boldsymbol{\sigma}\mathbf{q})_{\Omega} \equiv (\mathbf{q}, \boldsymbol{\sigma}\mathbf{n})_{\partial\Omega} \tag{4.19}$$

We decompose $\nabla\mathbf{q}$ into a symmetric part $(\nabla\mathbf{q})_s \equiv \frac{1}{2}\big(\nabla\mathbf{q} + (\nabla\mathbf{q})^T\big) \equiv \boldsymbol{\epsilon}(\mathbf{q})$ and skew-symmetric part $(\nabla\mathbf{q})_{ss}$ and note that the contraction between a symmetric and skew-symmetric tensor is zero to obtain

$$\boldsymbol{\sigma} : \nabla\mathbf{q} \equiv \boldsymbol{\sigma} : (\nabla\mathbf{q})_s + \cancelto{0}{\boldsymbol{\sigma} : (\nabla\mathbf{q})_{ss}} = \boldsymbol{\sigma} : \boldsymbol{\epsilon}(\mathbf{q}) \tag{4.20}$$

From (5.7), (5.8), (5.9) and (5.10), we get

$$(\boldsymbol{\sigma}\mathbf{n}, \mathbf{q})_{\partial\Omega} - (\boldsymbol{\sigma}, \boldsymbol{\epsilon}(\mathbf{q}))_{\Omega} + (\mathbf{f}, \mathbf{q})_{\Omega} = 0$$

which, after invoking the boundary condition $\boldsymbol{\sigma}\mathbf{n} = \mathbf{t}$ on Γ_N^p results in

$$(\mathbf{t}, \mathbf{q})_{\Gamma_N^p} - (\boldsymbol{\sigma}, \boldsymbol{\epsilon}(\mathbf{q}))_{\Omega} + (\mathbf{f}, \mathbf{q})_{\Omega} = 0 \tag{4.21}$$

The stress tensor $\boldsymbol{\sigma}$ and strain tensor $\boldsymbol{\epsilon}(\mathbf{q})$ are written as

$$\boldsymbol{\sigma} = \boldsymbol{s} + \frac{1}{3}tr(\boldsymbol{\sigma})\mathbf{I} = \boldsymbol{s} + \sigma_v\mathbf{I}; \qquad \boldsymbol{\epsilon}(\mathbf{q}) = \boldsymbol{e}(\mathbf{q}) + \frac{1}{3}tr(\boldsymbol{\epsilon}(\mathbf{q}))\mathbf{I} = \boldsymbol{e}(\mathbf{q}) + \frac{1}{3}\bar{\epsilon}(\mathbf{q})\mathbf{I}$$

where $\boldsymbol{s}$ is the deviatoric stress tensor, $\boldsymbol{e}(\mathbf{q})$ is the deviatoric strain tensor and σ_v is the mean stress. Using the above relations, we can write

$$\begin{aligned}
&\boldsymbol{\sigma} : \boldsymbol{\epsilon}(\mathbf{q}) = \big(\boldsymbol{s} + \sigma_v\mathbf{I}\big) : \big(\boldsymbol{e}(\mathbf{q}) + \frac{1}{3}\bar{\epsilon}(\mathbf{q})\mathbf{I}\big) \\
&= \boldsymbol{s} : \boldsymbol{e}(\mathbf{q}) + \boldsymbol{s} : \frac{1}{3}\bar{\epsilon}(\mathbf{q})\mathbf{I} + \sigma_v\mathbf{I} : \boldsymbol{e}(\mathbf{q}) + \sigma_v\mathbf{I} : \frac{1}{3}\bar{\epsilon}(\mathbf{q})\mathbf{I} \\
&= \boldsymbol{s} : \boldsymbol{e}(\mathbf{q}) + \frac{1}{3}\bar{\epsilon}(\mathbf{q})\cancelto{0}{tr(\boldsymbol{s})} + \sigma_v\cancelto{0}{tr(\boldsymbol{e}(\mathbf{q}))} + 3\sigma_v\frac{1}{3}\bar{\epsilon}(\mathbf{q}) = \boldsymbol{s} : \boldsymbol{e}(\mathbf{q}) + \sigma_v\bar{\epsilon}(\mathbf{q})
\end{aligned} \tag{4.22}$$

where we note that the contraction of any second order tensor with the identity tensor $\mathbf{I}$ is equal to the trace of the tensor and further, the trace of a deviatoric tensor is zero resulting in $tr(\boldsymbol{s}) = 0$ and $tr(\boldsymbol{e}(\mathbf{q})) = 0$. Substituting (4.22) in (4.21), we get

$$(\mathbf{t}, \mathbf{q})_{\Gamma_N^p} - (\mathbf{s}, \mathbf{e}(\mathbf{q}))_\Omega - (\sigma_v, \bar{\epsilon}(\mathbf{q}))_\Omega + (\mathbf{f}, \mathbf{q})_\Omega = 0 \tag{4.23}$$

The deviatoric strain tensor is obtained as

$$\begin{aligned} \boldsymbol{s} &= \boldsymbol{\sigma} - \frac{1}{3}tr(\boldsymbol{\sigma})\mathbf{I} = \lambda\bar{\epsilon}\mathbf{I} + 2G\boldsymbol{\epsilon} - \alpha p\mathbf{I} - \frac{1}{3}tr\big(\lambda\bar{\epsilon}\mathbf{I} + 2G\boldsymbol{\epsilon} - \alpha p\mathbf{I}\big)\mathbf{I} \\ &= 2G\big(\boldsymbol{\epsilon} - \frac{1}{3}tr(\boldsymbol{\epsilon})\mathbf{I}\big) = 2G\boldsymbol{e} \end{aligned} \tag{4.24}$$

Substituting (4.24) in (4.23), we get

$$(\mathbf{t}, \mathbf{q})_{\Gamma_N^p} - (2G\mathbf{e}, \mathbf{e}(\mathbf{q}))_\Omega - (\sigma_v, \bar{\epsilon}(\mathbf{q}))_\Omega + (\mathbf{f}, \mathbf{q})_\Omega = 0$$

As a result, the discrete variational statement of the linear momentum balance (2.6) for the m^{th} coupling iteration in the $(n+1)^{th}$ time step in the presence of medium heterogeneities is written as

$$\begin{aligned} &\sum_{E^p \in \mathscr{T}_H^p} 2G_{E^p}(\boldsymbol{e}(\mathbf{u}_H^{m,n+1}), \boldsymbol{e}(\mathbf{q}_H))_{E^p} + \sum_{E^p \in \mathscr{T}_H^p} (\sigma_v^{m,n+1}, \nabla \cdot \mathbf{q}_H)_{E^p} \\ &= \sum_{E^p \in \mathscr{T}_H^p} (\mathbf{f}, \mathbf{q}_H)_{E^p} + \sum_{E^p \in \mathscr{T}_H^p} (\mathbf{t}, \mathbf{q}_H)_{\partial E^p \cap \Gamma_N^p} \end{aligned} \tag{4.25}$$

Replacing m by $m+1$ in (4.25) and subtracting the two equations, we get

$$\sum_{E^p \in \mathscr{T}_H^p} 2G_{E^p}(\boldsymbol{e}(\delta^{(m)}\mathbf{u}_H), \boldsymbol{e}(\mathbf{q}_H))_{E^p} + \sum_{E^p \in \mathscr{T}_H^p} (\delta^{(m)}\sigma_v, \nabla \cdot \mathbf{q}_H)_{E^p} = 0$$

4.3.3 Summary of discrete variational statements in terms of coupling iteration differences

Find $\delta^{(m)}p_h \in W_h$, $\delta^{(m)}\mathbf{z}_h \in \mathbf{V}_h$ and $\delta^{(m)}\mathbf{u}_H \in \mathbf{U}_H$ such that

$$\begin{aligned}
&\sum_{E^f \in \mathscr{T}_h^f} \left(\frac{1}{M_{E^f}} + \frac{\alpha_{E^f}^2}{K_{b_{E^f}}}\right)(\delta^{(m)}p_h, \theta_h)_{E^f} + \sum_{E^f \in \mathscr{T}_h^f} \Delta t(\nabla \cdot \delta^{(m)}\mathbf{z}_h, \theta_h)_{E^f} \\
&= - \sum_{E^f \in \mathscr{T}_h^f} \frac{\alpha_{E^f}}{K_{b_{E^f}}}(\delta^{(m-1)}\sigma_v, \theta_h)_{E^f} \qquad (4.26)
\end{aligned}$$

$$\sum_{E^f \in \mathscr{T}_h^f} (\boldsymbol{\kappa}_{E^f}^{-1}\delta^{(m)}\mathbf{z}_h, \mathbf{v}_h)_{E^f} = \sum_{E^f \in \mathscr{T}_h^f} (\delta^{(m)}p_h, \nabla \cdot \mathbf{v}_h)_{E^f} \qquad (4.27)$$

$$\sum_{E^p \in \mathscr{T}_H^p} 2G_{E^p}(\boldsymbol{e}(\delta^{(m)}\mathbf{u}_H), \boldsymbol{e}(\mathbf{q}_H))_{E^p} + \sum_{E^p \in \mathscr{T}_H^p} \delta^{(m)}(\sigma_v, \nabla \cdot \mathbf{q}_H)_{E^p} = 0 \qquad (4.28)$$

where the finite dimensional spaces W_h, W_H, $\mathbf{V}_h$ and $\mathbf{U}_H$ are given by

$$W_h \equiv \left\{\theta_h : \theta_h|_{E^f} \in \mathbb{P}_0(E^f)\ \forall E^f \in \mathscr{T}_h^f\right\}$$

$$W_H \equiv \left\{\theta'_H : \theta'_H|_{E^p} \in \mathbb{P}_0(E^p)\ \forall E^p \in \mathscr{T}_H^p\right\}$$

$$\mathbf{V}_h \equiv \left\{\mathbf{v}_h : \mathbf{v}_h|_{E^f} \leftrightarrow \hat{\mathbf{v}}|_{\hat{E}} : \hat{\mathbf{v}}|_{\hat{E}} \in \hat{\mathbf{V}}(\hat{E})\ \forall E^f \in \mathscr{T}_h^f,\ \mathbf{v}_h \cdot \mathbf{n} = 0 \text{ on } \Gamma_N^f\right\}$$

$$\mathbf{U}_H \equiv \left\{\mathbf{q}_H = (u, v, w) : u|_{E^p}, v|_{E^p}, w|_{E^p} \in \mathbb{Q}_1(E^p)\ \forall E^p \in \mathscr{T}_H^p, \mathbf{q}_H = \mathbf{0} \text{ on } \Gamma_D^p\right\}$$

4.3.4 Restriction and prolongation operators

We introduce the restriction operator $\mathscr{R}$ that maps the fine scale pressure solution onto the coarse scale poromechanics grid and the prolongation operator $\mathscr{P}$ that maps the coarse scale volumetric strain onto the fine scale flow grid as follows

$$\mathscr{R} : W_h \mapsto W_H$$

$$\mathscr{P} : \nabla \cdot \mathbf{U}_H \mapsto W_h$$

As a result, the mean stress is defined on the fine and coarse grids as

$$\sigma_v = K_{b_{E^p}} \bar{\epsilon}_H - \alpha_{E^p} \mathscr{R} p_h \qquad (\forall\ E^p \in \mathscr{T}_H^p) \tag{4.29}$$

$$\sigma_v = K_{b_{E^f}} \mathscr{P} \bar{\epsilon}_H - \alpha_{E^p} p_h \qquad (\forall\ E^f \in \mathscr{T}_h^f) \tag{4.30}$$

Theorem 4.3.1. *In the presence of medium heterogeneities, the two-grid fixed stress split iterative scheme in which the flow subproblem is resolved on a finer grid is a contraction map given by*

$$\sum_{E^f \in \mathscr{T}_h^f} \frac{\|\delta^{(m)} \sigma_v\|_{E^f}^2}{K_{b_{E^f}}} + \overbrace{\sum_{E^p \in \mathscr{T}_H^p} 4G_{E^p} \|\boldsymbol{e}(\delta^{(m)} \mathbf{u}_H)\|_{E^p}^2 + \sum_{E^p \in \mathscr{T}_H^p} K_{b_{E^p}} \|\delta^{(m)} \bar{\epsilon}_H\|_{E^p}^2}^{>0}$$

$$+ \overbrace{\sum_{E^f \in \mathscr{T}_h^f} \Delta t\, 2 \|\boldsymbol{\kappa}_{E^f}^{-1/2} \delta^{(m)} \mathbf{z}_h\|_{E^f}^2}^{>0} \le \gamma \sum_{E^f \in \mathscr{T}_h^f} \frac{\|\delta^{(m-1)} \sigma_v\|_{E^f}^2}{K_{b_{E^f}}} \tag{4.31}$$

with contraction constant $\gamma \equiv \max\limits_{E^f \in \mathscr{T}_h^f} \left(\frac{\alpha_{E^f}^2}{\frac{K_{b_{E^f}}}{M_{E^f}} + \alpha_{E^f}^2} \right) < 1$, *if the following conditions are satisfied*

1. *First condition*

$$\sum_{E^f \in \mathscr{T}_h^f} (2\alpha_{E^f} \mathscr{P} \delta^{(m)}, \bar{\epsilon}_H \delta^{(m)} p_h)_{E^f} - \sum_{E^p \in \mathscr{T}_H^p} (2\alpha_{E^p} \delta^{(m)} \bar{\epsilon}_H, \mathscr{R} \delta^{(m)} p_h)_{E^p} = 0$$

2. *Second condition*

$$\sum_{E^p \in \mathscr{T}_H^p} K_{b_{E^p}} \|\delta^{(m)} \bar{\epsilon}_H\|_{E^p}^2 - \sum_{E^f \in \mathscr{T}_h^f} K_{b_{E^f}} \|\mathscr{P} \delta^{(m)} \bar{\epsilon}_H\|_{E^f}^2 = 0$$

Proof. • **Step 1: Flow equations**

Testing (4.26) with $\theta_h \in W_h$ such that $\theta_h|_{E^f} = \delta^{(m)} p_h \ \forall \ E^f \in \mathscr{T}_h^f$, we get

$$\begin{aligned} &\sum_{E^f \in \mathscr{T}_h^f} \left(\frac{1}{M_{E^f}} + \frac{\alpha_{E^f}^2}{K_{b_{E^f}}} \right) \|\delta^{(m)} p_h\|_{E^f}^2 + \sum_{E^f \in \mathscr{T}_h^f} \Delta t (\nabla \cdot \delta^{(m)} \mathbf{z}_h, \delta^{(m)} p_h)_{E^f} \\ &= - \sum_{E^f \in \mathscr{T}_h^f} \frac{\alpha_{E^f}}{K_{b_{E^f}}} (\delta^{(m-1)} \sigma_v, \delta^{(m)} p_h)_{E^f} \end{aligned} \tag{4.32}$$

Testing (4.27) with $\mathbf{v}_h \in \mathbf{V}_h$ such that $\mathbf{v}_h|_{E^f} \equiv \delta^{(m)} \mathbf{z}_h \ \forall \ E^f \in \mathscr{T}_h^f$, we get

$$\sum_{E^f \in \mathscr{T}_h^f} \|\boldsymbol{\kappa}_{E^f}^{-1/2} \delta^{(m)} \mathbf{z}_h\|_{E^f}^2 = \sum_{E^f \in \mathscr{T}_h^f} (\delta^{(m)} p_h, \nabla \cdot \delta^{(m)} \mathbf{z}_h)_{E^f} \tag{4.33}$$

From (4.32) and (4.33), we get

$$\begin{aligned} &\sum_{E^f \in \mathscr{T}_h^f} \left(\frac{1}{M_{E^f}} + \frac{\alpha_{E^f}^2}{K_{b_{E^f}}} \right) \|\delta^{(m)} p_h\|_{E^f}^2 + \sum_{E^f \in \mathscr{T}_h^f} \Delta t \|\boldsymbol{\kappa}_{E^f}^{-1/2} \delta^{(m)} \mathbf{z}_h\|_{E^f}^2 \\ &= \sum_{E^f \in \mathscr{T}_h^f} - \frac{\alpha_{E^f}}{K_{b_{E^f}}} (\delta^{(m-1)} \sigma_v, \delta^{(m)} p_h)_{E^f} \end{aligned} \tag{4.34}$$

• **Step 2: Invoking the Young's inequality**

Since the terms on the LHS of (4.34) are strictly positive, the RHS is also strictly positive. We invoke the Young's inequality for products for the RHS of (4.34) as follows

$$\begin{aligned} &- \frac{\alpha_{E^f}}{K_{b_{E^f}}} \delta^{(m-1)} \sigma_v, \delta^{(m)} p_h \leq \frac{1}{2 \varepsilon_{E^f} K_{b_{E^f}}^2} |\delta^{(m-1)} \sigma_v|^2 + \frac{\varepsilon_{E^f}}{2} |\alpha_{E^f} \delta^{(m)} p_h|^2 \\ &\quad (\forall \ E^f \in \mathscr{T}_h^f) \end{aligned}$$

Since the above inequality is true for any $\varepsilon_{E^f} > 0$, we choose

$$\varepsilon_{E^f} = \frac{1}{\alpha_{E^f}^2} \left(\frac{1}{M_{E^f}} + \frac{\alpha_{E^f}^2}{K_{b_{E^f}}} \right)$$

to get the following

$$
\begin{aligned}
&-\frac{\alpha_{E^f}}{K_{b_{E^f}}}(\delta^{(m-1)}\sigma_v, \delta^{(m)}p_h)_{E^f} \\
&\leq \frac{\alpha^2_{E^f}}{2K_{b_{E^f}}\left(\frac{K_{b_{E^f}}}{M_{E^f}} + \alpha^2_{E^f}\right)}\|\delta^{(m-1)}\sigma_v\|^2_{E^f} + \frac{1}{2}\left(\frac{1}{M_{E^f}} + \frac{\alpha^2_{E^f}}{K_{b_{E^f}}}\right)\|\delta^{(m)}p_h\|^2_{E^f} \\
&\quad (\forall\ E^f \in \mathscr{T}^f_h)
\end{aligned}
\tag{4.35}
$$

From (4.34) and (4.35), we get

$$
\begin{aligned}
&\sum_{E^f\in\mathscr{T}^f_h}\left(\frac{1}{M_{E^f}} + \frac{\alpha^2_{E^f}}{K_{b_{E^f}}}\right)\|\delta^{(m)}p_h\|^2_{E^f} + \sum_{E^f\in\mathscr{T}^f_h}\Delta t\|\boldsymbol{\kappa}^{-1/2}_{E^f}\delta^{(m)}\mathbf{z}_h\|^2_{E^f} \\
&\leq \sum_{E^f\in\mathscr{T}^f_h}\frac{\alpha^2_{E^f}}{2K_{b_{E^f}}\left(\frac{K_{b_{E^f}}}{M_{E^f}} + \alpha^2_{E^f}\right)}\|\delta^{(m-1)}\sigma_v\|^2_{E^f} \\
&+ \sum_{E^f\in\mathscr{T}^f_h}\frac{1}{2}\left(\frac{1}{M_{E^f}} + \frac{\alpha^2_{E^f}}{K_{b_{E^f}}}\right)\|\delta^{(m)}p_h\|^2_{E^f}
\end{aligned}
$$

which can also be written as

$$
\begin{aligned}
&\sum_{E^f\in\mathscr{T}^f_h}\frac{1}{2}\left(\frac{1}{M_{E^f}} + \frac{\alpha^2_{E^f}}{K_{b_{E^f}}}\right)\|\delta^{(m)}p_h\|^2_{E^f} + \sum_{E^f\in\mathscr{T}^f_h}\Delta t\|\boldsymbol{\kappa}^{-1/2}_{E^f}\delta^{(m)}\mathbf{z}_h\|^2_{E^f} \\
&\leq \sum_{E^f\in\mathscr{T}^f_h}\frac{\alpha^2_{E^f}}{2K_{b_{E^f}}\left(\frac{K_{b_{E^f}}}{M_{E^f}} + \alpha^2_{E^f}\right)}\|\delta^{(m-1)}\sigma_v\|^2_{E^f}
\end{aligned}
$$

which, after noting that $\left(\frac{1}{M_{E^f}} + \frac{\alpha^2_{E^f}}{K_{b_{E^f}}}\right) > \frac{\alpha^2_{E^f}}{K_{b_{E^f}}}$, can also be written as

$$
\begin{aligned}
&\sum_{E^f\in\mathscr{T}^f_h}\frac{\alpha^2_{E^f}}{K_{b_{E^f}}}\|\delta^{(m)}p_h\|^2_{E^f} + \sum_{E^f\in\mathscr{T}^f_h}\Delta t\, 2\|\boldsymbol{\kappa}^{-1/2}_{E^f}\delta^{(m)}\mathbf{z}_h\|^2_{E^f} \\
&\leq \sum_{E^f\in\mathscr{T}^f_h}\frac{\alpha^2_{E^f}}{K_{b_{E^f}}\left(\frac{K_{b_{E^f}}}{M_{E^f}} + \alpha^2_{E^f}\right)}\|\delta^{(m-1)}\sigma_v\|^2_{E^f}
\end{aligned}
\tag{4.36}
$$

• **Step 3: Poromechanics equations**

Testing (4.28) with $\mathbf{q}_H \in \mathbf{Q}_H$ such that $\mathbf{q}|_{E^p} = 2\delta^{(m)}\mathbf{u}_H \; \forall \; E^p \in \mathscr{T}_H^p$ and noting that $\nabla \cdot \delta^{(m)}\mathbf{u}_H \equiv \delta^{(m)}\bar{\epsilon}_H$, we get

$$\sum_{E^p \in \mathscr{T}_H^p} 4G_{E^p} \|\boldsymbol{e}(\delta^{(m)}\mathbf{u}_H)\|^2_{E^p} + \sum_{E^p \in \mathscr{T}_H^p} 2(\delta^{(m)}\sigma_v, \delta^{(m)}\bar{\epsilon}_H)_{E^p} = 0 \tag{4.37}$$

Further, from (4.29), we note that $\delta^{(m)}\sigma_v = K_{b_{E^p}}\delta^{(m)}\bar{\epsilon}_H - \alpha_{E^p}\mathscr{R}\delta^{(m)}p_h \; \forall \; E^p \in \mathscr{T}_H^p$. As a result, (4.37) is written as

$$\begin{aligned}&\sum_{E^p \in \mathscr{T}_H^p} 4G_{E^p} \|\boldsymbol{e}(\delta^{(m)}\mathbf{u}_H)\|^2_{E^p} + \sum_{E^p \in \mathscr{T}_H^p} 2K_{b_{E^p}} \|\delta^{(m)}\bar{\epsilon}_H\|^2_{E^p} \\ &- \sum_{E^p \in \mathscr{T}_H^p} 2\alpha_{E^p}(\delta^{(m)}\bar{\epsilon}_H, \mathscr{R}\delta^{(m)}p_h)_{E^p} = 0\end{aligned} \tag{4.38}$$

• **Step 4: Combining flow and poromechanics equations**

Adding (4.36) and (4.38), we get

$$\begin{aligned}&\sum_{E^f \in \mathscr{T}_h^f} \frac{\alpha^2_{E^f}}{K_{b_{E^f}}} \|\delta^{(m)}p_h\|^2_{E^f} + \sum_{E^f \in \mathscr{T}_h^f} \Delta t\, 2\|\boldsymbol{\kappa}_{E^f}^{-1/2}\delta^{(m)}\mathbf{z}_h\|^2_{E^f} \\ &+ \sum_{E^p \in \mathscr{T}_H^p} 4G_{E^p} \|\boldsymbol{e}(\delta^{(m)}\mathbf{u}_H)\|^2_{E^p} + \sum_{E^p \in \mathscr{T}_H^p} 2K_{b_{E^p}} \|\delta^{(m)}\bar{\epsilon}_H\|^2_{E^p} \\ &- \sum_{E^p \in \mathscr{T}_H^p} 2\alpha_{E^p}(\delta^{(m)}\bar{\epsilon}_H, \mathscr{R}\delta^{(m)}p_h)_{E^p} \leq \sum_{E^f \in \mathscr{T}_h^f} \frac{\alpha^2_{E^f}}{\left(\frac{K_{b_{E^f}}}{M_{E^f}} + \alpha^2_{E^f}\right)} \frac{\|\delta^{(m-1)}\sigma_v\|^2_{E^f}}{K_{b_{E^f}}}\end{aligned} \tag{4.39}$$

Now, from (4.30), we note that

$$\begin{aligned}\|\delta^{(m)}\sigma_v\|^2_{E^f} &= \alpha^2_{E^f}\|\delta^{(m)}p_h\|^2_{E^f} + K^2_{b_{E^f}}\|\mathscr{P}\delta^{(m)}\bar{\epsilon}_H\|^2_{E^f} \\ &\quad - 2K_{b_{E^f}}\alpha_{E^f}(\mathscr{P}\delta^{(m)}\bar{\epsilon}_H, \delta^{(m)}p_h)_{E^f} \qquad (\forall \; E^f \in \mathscr{T}_h^f)\end{aligned}$$

which implies that

$$\frac{\alpha_{E^f}^2}{K_{b_{E^f}}}\|\delta^{(m)}p_h\|_{E^f}^2 = \frac{\|\delta^{(m)}\sigma_v\|_{E^f}^2}{K_{b_{E^f}}} - K_{b_{E^f}}\|\mathscr{P}\delta^{(m)}\bar{\epsilon}_H\|_{E^f}^2$$
$$+ 2\alpha_{E^f}(\mathscr{P}\delta^{(m)}\bar{\epsilon}_H, \delta^{(m)}p_h)_{E^f} \qquad (\forall\ E^f \in \mathscr{T}_h^f) \tag{4.40}$$

Substituting (4.40) in (4.39), we get

$$\sum_{E^f\in\mathscr{T}_h^f}\frac{\|\delta^{(m)}\sigma_v\|_{E^f}^2}{K_{b_{E^f}}} + \overbrace{\sum_{E^p\in\mathscr{T}_H^p} 4G_{E^p}\|\boldsymbol{e}(\delta^{(m)}\mathbf{u}_H)\|_{E^p}^2 + \sum_{E^p\in\mathscr{T}_H^p} K_{b_{E^p}}\|\delta^{(m)}\bar{\epsilon}_H\|_{E^p}^2}^{>0}$$
$$+\left[\overbrace{\sum_{E^f\in\mathscr{T}_h^f} 2\alpha_{E^f}(\mathscr{P}\delta^{(m)}\bar{\epsilon}_H, \delta^{(m)}p_h)_{E^f} - \sum_{E^p\in\mathscr{T}_H^p} 2\alpha_{E^p}(\delta^{(m)}\bar{\epsilon}_H, \mathscr{R}\delta^{(m)}p_h)_{E^p}}^{\text{Set} = 0 \text{ to obtain expressions for } \mathrm{K}_{\mathrm{b}_{\mathrm{E^p}}}\ \forall\ \mathrm{E^p}\in\mathscr{T}_{\mathrm{H}}^{\mathrm{p}} \text{ and } \mathscr{P}\delta^{(\mathrm{m})}\bar{\epsilon}_{\mathrm{H}}\ \forall\ \mathrm{E^f}\in\mathscr{T}_{\mathrm{h}}^{\mathrm{f}}}\right]$$
$$+\left[\overbrace{\sum_{E^p\in\mathscr{T}_H^p} K_{b_{E^p}}\|\delta^{(m)}\bar{\epsilon}_H\|_{E^p}^2 - \sum_{E^f\in\mathscr{T}_h^f} K_{b_{E^f}}\|\mathscr{P}\delta^{(m)}\bar{\epsilon}_H\|_{E^f}^2}^{\text{Turns out to be } \geq 0 \text{ in lieu of Cauchy}-\text{Schwartz inequality}}\right]$$
$$+\overbrace{\sum_{E^f\in\mathscr{T}_h^f}\Delta t\, 2\|\boldsymbol{\kappa}_{E^f}^{-1/2}\delta^{(m)}\mathbf{z}_h\|_{E^f}^2}^{>0} \leq \gamma \sum_{E^f\in\mathscr{T}_h^f}\frac{\|\delta^{(m-1)}\sigma_v\|_{E^f}^2}{K_{b_{E^f}}} \tag{4.41}$$

The statement (4.41) is a contraction map in a sense that

$$\sum_{E^f\in\mathscr{T}_h^f}\frac{\|\delta^{(0)}\sigma_v\|_{E^f}^2}{K_{b_{E^f}}} > \sum_{E^f\in\mathscr{T}_h^f}\frac{\|\delta^{(1)}\sigma_v\|_{E^f}^2}{K_{b_{E^f}}} > \sum_{E^f\in\mathscr{T}_h^f}\frac{\|\delta^{(2)}\sigma_v\|_{E^f}^2}{K_{b_{E^f}}} > \ldots$$

with contraction constant γ given by

$$\gamma \equiv \max_{E^f\in\mathscr{T}_h^f}\frac{\alpha_{E^f}^2}{\left(\frac{K_{b_{E^f}}}{M_{E^f}} + \alpha_{E^f}^2\right)} < 1$$

provided the following are true

$$\sum_{E^f\in\mathscr{T}_h^f} 2\alpha_{E^f}(\mathscr{P}\delta^{(m)}\bar{\epsilon}_H, \delta^{(m)}p_h)_{E^f} - \sum_{E^p\in\mathscr{T}_H^p} 2\alpha_{E^p}(\delta^{(m)}\bar{\epsilon}_H, \mathscr{R}\delta^{(m)}p_h)_{E^p} = 0 \tag{4.42}$$

$$\sum_{E^p \in \mathscr{T}_H^p} K_{b_{E^p}} \|\delta^{(m)} \bar{\epsilon}_H\|_{E^p}^2 - \sum_{E^f \in \mathscr{T}_h^f} K_{b_{E^f}} \|\mathscr{P} \delta^{(m)} \bar{\epsilon}_H\|_{E^f}^2 \geq 0 \tag{4.43}$$

□

The objective now is to satisfy the conditions (4.42) and (4.43) for the convergence of the two-grid fixed stress split scheme.

4.4 Satisfaction of conditions for convergence of the fully discrete two-grid fixed stress split scheme

Corollary 4.4.0.1. *Satisfaction of the fixed mean stress constraint during the flow solve at both scales leads to the following expressions for the upscaled pore pressures*

$$\mathscr{R} \delta^{(m)} p_h = \frac{K_{b_{E^p}}}{\alpha_{E^p} Meas(E^p)} \sum_{E^f \in \mathfrak{I}^{E^p}} \frac{\alpha_{E^f}}{K_{b_{E^f}}} \delta^{(m)} p_h Meas(E^f) \qquad (\forall\ E^p \in \mathfrak{T}_H^p)$$

Proof. • **Step 1: Using the fact that pore pressure is frozen during the poromechanical solve**

Since the pore pressure is frozen during the poromechanical solve, the total pore pressure change in a coupling iteration is the same as the pore pressure change calculated during the flow solve in the coupling iteration as follows

$$\mathscr{R} \delta_f^{(m)} p_h = \mathscr{R} \delta^{(m)} p_h \qquad (\forall\ E^p \in \mathscr{T}_H^p) \tag{4.44}$$

$$\delta_f^{(m)} p_h = \delta^{(m)} p_h \qquad (\forall\ E^f \in \mathscr{T}_h^f) \tag{4.45}$$

• **Step 2: Applying the fixed mean stress constraint on both scales**

Now, the fixed mean stress constraint implies that there is no change in the mean stress of the system during the flow solve. This naturally implies that

$$\int_{E^p} \delta_f^{(m)} \sigma_v = 0 \qquad (\forall\ E^p \in \mathscr{T}_H^p)$$

In lieu of (4.29), we write the above as

$$\int_{E^p} (K_{b_{E^p}} \delta_f^{(m)} \bar{\epsilon}_H - \alpha_{E^p} \mathscr{R} \delta_f^{(m)} p_h) = 0 \qquad (\forall\ E^p \in \mathscr{T}_H^p)$$

which, in lieu of (4.44), can be written as

$$\int_{E^p} \delta_f^{(m)} \bar{\epsilon}_H = \frac{\alpha_{E^p}}{K_{b_{E^p}}} \mathscr{R} \delta^{(m)} p_h Meas(E^p) \qquad (\forall\ E^p \in \mathscr{T}_H^p) \tag{4.46}$$

Denoting $\int_{E^f} \delta_f^{(m)} \mathscr{P} \bar{\epsilon}_H$ is the change in volume of each element E^f of $\mathfrak{J}^{E^p}$, we now impose the fixed mean stress constraint on each element E^f of $\mathfrak{J}^{E^p}$ as follows

$$\int_{E^f} \delta_f^{(m)} \sigma_v \equiv \int_{E^f} (K_{b_{E^f}} \mathscr{P} \delta_f^{(m)} \bar{\epsilon}_H - \alpha_{E^f} \delta^{(m)} p_h) = 0 \qquad (\forall\ E^f \in \mathfrak{J}^{E^p})$$

which, in lieu of (4.45), can be written as

$$\int_{E^f} \mathscr{P} \delta_f^{(m)} \bar{\epsilon}_H = \frac{\alpha_{E^f}}{K_{b_{E^f}}} \delta^{(m)} p_h Meas(E^f) \qquad (\forall\ E^f \in \mathscr{T}_h^f) \tag{4.47}$$

• **Step 3: Applying the volume balance criterion to coarse scale poromechanical element(s)**

The term $\int_{E^p} \delta_f^{(m)} \bar{\epsilon}_H$ is the change in volume of E^p during the flow solve in the $(m+1)^{th}$ coupling iteration. This naturally equates the sum of corresponding changes in volumes of the elements of $\mathfrak{J}^{E^p}$ as follows

$$\int_{E^p} \delta_f^{(m)} \bar{\epsilon}_H \equiv \sum_{E^f \in \mathfrak{J}^{E^p}} \int_{E^f} \mathscr{P} \delta_f^{(m)} \bar{\epsilon}_H \qquad (\forall\ E^p \in \mathscr{T}_H^p) \tag{4.48}$$

From (4.47) and (4.48), we get

$$\int\limits_{E^p} \delta_f^{(m)} \bar{\epsilon}_H = \sum_{E^f \in \mathfrak{I}^{E^p}} \frac{\alpha_{E^f}}{K_{b_{E^f}}} \delta^{(m)} p_h Meas(E^f) \qquad (\forall\ E^p \in \mathscr{T}_H^p) \tag{4.49}$$

From (4.46) and (4.49), we get

$$\sum_{E^f \in \mathfrak{I}^{E^p}} \frac{\alpha_{E^f}}{K_{b_{E^f}}} \delta^{(m)} p_h Meas(E^f) = \frac{\alpha_{E^p}}{K_{b_{E^p}}} \mathscr{R} \delta^{(m)} p_h Meas(E^p) \qquad (\forall\ E^p \in \mathscr{T}_H^p)$$

which results in

$$\mathscr{R} \delta^{(m)} p_h = \frac{K_{b_{E^p}}}{\alpha_{E^p} Meas(E^p)} \sum_{E^f \in \mathfrak{I}^{E^p}} \frac{\alpha_{E^f}}{K_{b_{E^f}}} \delta^{(m)} p_h Meas(E^f) \qquad (\forall\ E^p \in \mathscr{T}_H^p) \tag{4.50}$$

□

Corollary 4.4.0.2. *Satisfaction of the condition* (4.42) *leads to the following expressions for the effective bulk moduli for the coarse scale poromechanical solve*

$$K_{b_{E^p}} = \frac{1}{\sum\limits_{E^f \in \mathfrak{I}^{E^p}} \frac{1}{K_{b_{E^f}}} \frac{Meas(E^f)}{Meas(E^p)}} \qquad (\forall\ E^p \in \mathscr{T}_H^p)$$

and the following expressions for the downscaled volumetric strains

$$\mathscr{P} \delta^{(m)} \bar{\epsilon}_H = \frac{K_{b_{E^p}}}{K_{b_{E^f}}} \frac{1}{Meas(E^p)} \int\limits_{E^p} \delta^{(m)} \bar{\epsilon}_H \qquad (\forall\ E^f \in \mathfrak{I}^{E^p}\ \forall\ E^p \in \mathscr{T}_H^p)$$

Proof. • **Step 1: Recasting the first term on LHS of** (4.42)

We start by modifying the first term on LHS of (4.42) as follows

$$\sum_{E^f \in \mathscr{T}_h^f} 2\alpha_{E^f} (\mathscr{P} \delta^{(m)} \bar{\epsilon}_H, \delta^{(m)} p_h)_{E^f} = \sum_{E^f \in \mathscr{T}_h^f} 2\alpha_{E^f} \delta^{(m)} p_h \mathscr{P} \delta^{(m)} \bar{\epsilon}_H Meas(E^f) \tag{4.51}$$

where we note that $\delta^{(m)}p_h \in W_h$. Since a flow element $E^f \in \mathscr{T}_h^f$ in uniquely associated with a poromechanical element E^p via $\mathfrak{I}^{E^p}$, we can write

$$\begin{aligned}&\sum_{E^f \in \mathscr{T}_h^f} 2\alpha_{E^f}\delta^{(m)}p_h \mathscr{P}\delta^{(m)}\bar{\epsilon}_H Meas(E^f)\\ &= \sum_{E^p \in \mathscr{T}_H^p} \sum_{E^f \in \mathfrak{I}^{E^p}} 2\alpha_{E^f}\delta^{(m)}p_h \mathscr{P}\delta^{(m)}\bar{\epsilon}_H Meas(E^f)\end{aligned} \tag{4.52}$$

In lieu of (4.51) and (4.52), we can write

$$\begin{aligned}&\sum_{E^f \in \mathscr{T}_h^f} 2\alpha_{E^f}(\mathscr{P}\delta^{(m)}\bar{\epsilon}_H, \delta^{(m)}p_h)_{E^f}\\ &= \sum_{E^p \in \mathscr{T}_H^p} \sum_{E^f \in \mathfrak{I}^{E^p}} 2\alpha_{E^f}\delta^{(m)}p_h \mathscr{P}\delta^{(m)}\bar{\epsilon}_H Meas(E^f)\end{aligned} \tag{4.53}$$

- **Step 2: Recasting the second term on LHS of** (4.42)

Next, we modify the second term on LHS of (4.42) as follows

$$\sum_{E^p \in \mathscr{T}_H^p} 2\alpha_{E^p}(\delta^{(m)}\bar{\epsilon}_H, \mathscr{R}\delta^{(m)}p_h)_{E^p} \equiv \sum_{E^p \in \mathscr{T}_H^p} 2\alpha_{E^p}\mathscr{R}\delta^{(m)}p_h \int_{E^p} \delta^{(m)}\bar{\epsilon}_H \tag{4.54}$$

where we note that $\mathscr{R}\delta^{(m)}p_h \in W_H$. In lieu of (4.53) and (4.54), the first condition given by (4.42) is rewritten as

$$\begin{aligned}&\sum_{E^p \in \mathscr{T}_H^p} \sum_{E^f \in \mathfrak{I}^{E^p}} 2\alpha_{E^f}\delta^{(m)}p_h \mathscr{P}\delta^{(m)}\bar{\epsilon}_H Meas(E^f)\\ &= \sum_{E^p \in \mathscr{T}_H^p} 2\alpha_{E^p}\mathscr{R}\delta^{(m)}p_h \int_{E^p} \delta^{(m)}\bar{\epsilon}_H\end{aligned} \tag{4.55}$$

- **Step 3: Substituting the expression for upscaled pore pressures**

Substituting the expression (4.50) in (4.55), we get

$$\sum_{E^p \in \mathscr{T}_H^p} \sum_{E^f \in \mathfrak{I}^{E^p}} 2\alpha_{E^f}\delta^{(m)}p_h \mathscr{P}\delta^{(m)}\bar{\epsilon}_H Meas(E^f)$$

$$= \sum_{E^p \in \mathscr{T}_H^p} 2\alpha_{E^p} \overbrace{\frac{K_{b_{E^p}}}{\alpha_{E^p} Meas(E^p)} \sum_{E^f \in \mathcal{J}^{E^p}} \frac{\alpha_{E^f}}{K_{b_{E^f}}} \delta^{(m)} p_h Meas(E^f)}^{\mathscr{R}\delta^{(m)} p_h} \int\limits_{E^p} \delta^{(m)} \bar{\epsilon}_H$$

which implies that

$$\sum_{E^p \in \mathscr{T}_H^p} \sum_{E^f \in \mathcal{J}^{E^p}} \left(\mathscr{P}\delta^{(m)}\bar{\epsilon}_H - \frac{K_{b_{E^p}}}{K_{b_{E^f}}} \frac{1}{Meas(E^p)} \int\limits_{E^p} \delta^{(m)}\bar{\epsilon}_H \right) \alpha_{E^f} \delta^{(m)} p_h Meas(E^f) = 0$$

which, in lieu of the linear independence of the basis p_h ($\forall\ E^f \in \mathscr{T}_h^f$) of the pressure space on the fine scale flow grid, implies that

$$\mathscr{P}\delta^{(m)}\bar{\epsilon}_H - \frac{K_{b_{E^p}}}{K_{b_{E^f}}} \frac{1}{Meas(E^p)} \int\limits_{E^p} \delta^{(m)}\bar{\epsilon}_H = 0 \qquad (\forall\ E^f \in \mathcal{J}^{E^p}\ \forall\ E^p \in \mathscr{T}_H^p)$$

implying that

$$\mathscr{P}\delta^{(m)}\bar{\epsilon}_H = \frac{K_{b_{E^p}}}{K_{b_{E^f}}} \frac{1}{Meas(E^p)} \int\limits_{E^p} \delta^{(m)}\bar{\epsilon}_H \qquad (\forall\ E^f \in \mathcal{J}^{E^p}\ \forall\ E^p \in \mathscr{T}_H^p) \quad (4.56)$$

- **Step 4: Applying the volume balance criterion to coarse scale poromechanical element(s)**

The change in volume of E^p over the $(m+1)^{th}$ coupling iteration equates the sum of corresponding changes in volumes of the elements of $\mathcal{J}^{E^p}$ as follows

$$\int\limits_{E^p} \delta^{(m)}\bar{\epsilon}_H = \sum_{E^f \in \mathcal{J}^{E^p}} \int\limits_{E^f} \mathscr{P}\delta^{(m)}\bar{\epsilon}_H = \sum_{E^f \in \mathcal{J}^{E^p}} \mathscr{P}\delta^{(m)}\bar{\epsilon}_H Meas(E^f) \qquad (\forall\ E^p \in \mathscr{T}_H^p) \quad (4.57)$$

In lieu of (4.56) and (4.57), we get

$$\int\limits_{E^p} \delta^{(m)}\bar{\epsilon}_H = \int\limits_{E^p} \delta^{(m)}\bar{\epsilon}_H \sum_{E^f \in \mathcal{J}^{E^p}} \frac{K_{b_{E^p}}}{K_{b_{E^f}}} \frac{Meas(E^f)}{Meas(E^p)} \qquad (\forall\ E^p \in \mathscr{T}_H^p)$$

which finally leads to

$$\frac{1}{K_{b_{E^p}}} = \sum_{E^f \in \mathfrak{J}^{E^p}} \frac{1}{K_{b_{E^f}}} \frac{Meas(E^f)}{Meas(E^p)} \qquad (\forall\ E^p \in \mathscr{T}_H^p) \tag{4.58}$$

□

Corollary 4.4.0.3. *The Cauchy-Schwartz inequality, along with the obtained expressions for effective coarse scale bulk moduli* (4.58) *and downscaled volumetric strains* (4.56)*, guarantees the satisfaction of the condition* (4.43).

Proof. • **Step 1: Recasting** (4.43) **in lieu of** (4.56) **and** (4.58)

The condition (4.43) given by

$$\sum_{E^p \in \mathscr{T}_H^p} K_{b_{E^p}} \|\delta^{(m)} \bar{\epsilon}_H\|_{E^p}^2 - \sum_{E^f \in \mathscr{T}_h^f} K_{b_{E^f}} \|\mathscr{P} \delta^{(m)} \bar{\epsilon}_H\|_{E^f}^2 \geq 0$$

can be written as

$$\sum_{E^p \in \mathscr{T}_H^p} K_{b_{E^p}} \|\delta^{(m)} \bar{\epsilon}_H\|_{E^p}^2 - \sum_{E^f \in \mathscr{T}_H^f} K_{b_{E^f}} |\mathscr{P} \delta^{(m)} \bar{\epsilon}_H|^2 Meas(E^f) \geq 0$$

which can also be written as

$$\sum_{E^p \in \mathscr{T}_H^p} \left[K_{b_{E^p}} \|\delta^{(m)} \bar{\epsilon}_H\|_{E^p}^2 - \sum_{E^f \in \mathfrak{J}^{E^p}} K_{b_{E^f}} |\mathscr{P} \delta^{(m)} \bar{\epsilon}_H|^2 Meas(E^f) \right] \geq 0$$

which, in lieu of (4.56), can also be written as

$$\sum_{E^p \in \mathscr{T}_H^p} \left[K_{b_{E^p}} \|\delta^{(m)} \bar{\epsilon}_H\|_{E^p}^2 - \sum_{E^f \in \mathfrak{J}^{E^p}} K_{b_{E^f}} \overbrace{\left(\frac{K_{b_{E^p}}}{K_{b_{E^f}}} \frac{1}{Meas(E^p)} \int_{E^p} \delta^{(m)} \bar{\epsilon}_H \right)^2}^{\|\mathscr{P} \delta^{(m)} \bar{\epsilon}_H\|^2} Meas(E^f) \right] \geq 0$$

which can also be written as

$$\sum_{E^p \in \mathscr{T}_H^p} \left[K_{b_{E^p}} \|\delta^{(m)} \bar{\epsilon}_H\|^2_{E^p} - \frac{K_{b_{E^p}}}{Meas(E^p)} \left(\int\limits_{E^p} |\delta^{(m)} \bar{\epsilon}_H| \right)^2 K_{b_{E^p}} \sum_{E^f \in \mathbb{J}^{E^p}} \frac{1}{K_{b_{E^f}}} \frac{Meas(E^f)}{Meas(E^p)} \right] \geq 0$$

which, in lieu of (4.58), can be written as

$$\sum_{E^p \in \mathscr{T}_H^p} \left[K_{b_{E^p}} \|\delta^{(m)} \bar{\epsilon}_H\|^2_{E^p} - \frac{K_{b_{E^p}}}{Meas(E^p)} \left(\int\limits_{E^p} |\delta^{(m)} \bar{\epsilon}_H| \right)^2 K_{b_{E^p}} \frac{1}{K_{b_{E^p}}} \right] \geq 0$$

which can be finally written as

$$\sum_{E^p \in \mathscr{T}_H^p} \left[K_{b_{E^p}} \|\delta^{(m)} \bar{\epsilon}_H\|^2_{E^p} - \frac{K_{b_{E^p}}}{Meas(E^p)} \left(\int\limits_{E^p} |\delta^{(m)} \bar{\epsilon}_H| \right)^2 \right] \geq 0 \tag{4.59}$$

- **Step 2: Applying the Cauchy-Schwartz inequality**

The Cauchy-Schwartz inequality (see Oden and Demkowicz [63]) states that if S is a measurable subset of $\mathbb{R}^3$, and f and g are measurable real- or complex-valued functions on S, then

$$\left(\int\limits_S |fg| \right)^2 \leq \int\limits_S |f|^2 \int\limits_S |g|^2$$

Replacing S by E^p, f by $\delta^{(m)} \bar{\epsilon}_H$ and g by 1, we get

$$\int\limits_{E^p} |\delta^{(m)} \bar{\epsilon}_H|^2 \geq \frac{1}{Meas(E^p)} \left(\int\limits_{E^p} |\delta^{(m)} \bar{\epsilon}_H| \right)^2 \qquad (\forall \ E^p \in \mathscr{T}_H^p)$$

which can be written as

$$K_{b_{E^p}} \|\delta^{(m)} \bar{\epsilon}_H\|^2_{E^p} - \frac{K_{b_{E^p}}}{Meas(E^p)} \left(\int\limits_{E^p} |\delta^{(m)} \bar{\epsilon}_H| \right)^2 \geq 0 \qquad (\forall \ E^p \in \mathscr{T}_H^p)$$

which implies that

$$\sum_{E^p \in \mathscr{T}_H^p} \left[K_{b_{E^p}} \|\delta^{(m)} \bar{\epsilon}_H\|^2_{E^p} - \frac{K_{b_{E^p}}}{Meas(E^p)} \left(\int\limits_{E^p} |\delta^{(m)} \bar{\epsilon}_H| \right)^2 \right] \geq 0$$

which is identical to (4.59). Thus, provided the downscaled volumetric strains are computed in accordance with (4.56) and effective coarse scale bulk moduli are computed in accordance with (4.58), the Cauchy-Schwartz inequality guarantees the satisfaction of the condition (4.43). □

4.5 The two-grid fixed stress split algorithm

Algorithm 5 Solution phase of two-grid fixed-stress split iterative scheme for consolidation in heterogeneous porous medium

```
For each time $t^n$
repeat                                                    ▷ Fixed-stress iteration # $(m+1)$
    Solve the flow equations for $p_h^{m+1,n+1}$ and $\mathbf{z}_h^{m+1,n+1}$
    Compute $\delta^{(m)} p_h \equiv p_h^{m+1,n+1} - p_h^{m,n+1}$
    Obtain $\mathscr{R}\delta^{(m)} p_h$                                 ▷ Upscale pore pressures
    Obtain $\mathscr{R} p_h^{m+1,n+1} \equiv \mathscr{R} p_h^{m,n+1} + \mathscr{R}\delta^{(m)} p_h$
    Using $\mathscr{R} p_h^{m+1,n+1}$, solve the poromechanics equations for $\mathbf{u}_H^{m+1,n+1}$
    Compute $\delta^{(m)}\bar{\epsilon}_H \equiv \delta^{(m)}\nabla\cdot\mathbf{u}_H = \nabla\cdot\mathbf{u}_H^{m+1,n+1} - \nabla\cdot\mathbf{u}_H^{m,n+1}$
    Obtain $\mathscr{P}\delta^{(m)}\bar{\epsilon}_H$                          ▷ Downscale volumetric strains
    Obtain $\mathscr{P}\bar{\epsilon}_H^{m+1,n+1} \equiv \mathscr{P}\delta^{(m)}\bar{\epsilon}_H + \mathscr{P}\bar{\epsilon}_H^{m,n+1}$
    Obtain $\delta^{(m)}\sigma_v = K_{b_{E^f}} \mathscr{P}\delta^{(m)}\bar{\epsilon}_H - \alpha_{E^f}\delta^{(m)} p_h$
until Stopping criterion

$$\frac{\displaystyle\sum_{E^f \in \mathscr{T}_h^f} \frac{\|\delta^{(m)}\sigma_v\|_{E^f}^2}{K_{b_{E^f}}}}{\displaystyle\sum_{E^f \in \mathscr{T}_h^f} \frac{\|\sigma_v^{m+1,n+1}\|_{E^f}^2}{K_{b_{E^f}}}} < TOL, \qquad TOL > 0$$

Increment $t^n \to t^{n+1}$
```

In the pre-processing step, the effective bulk modulus for coarse scale poromechanical solve is obtained in terms of fine scale bulk modulus as

$$K_{b_{E^p}} = \frac{1}{\sum\limits_{E^f \in \mathfrak{J}^{E^p}} \frac{1}{K_{b_{E^f}}} \frac{Meas(E^f)}{Meas(E^p)}} \qquad (\forall\ E^p \in \mathscr{T}_H^p)$$

As presented in Algorithm 5, the mapping operators alternatively upscale the pore pressures and downscale the volumetric strains in every fixed stress iteration in each time step in accordance with

$$\mathscr{R}\delta^{(m)}p_h = \frac{K_{b_{E^p}}}{\alpha_{E^p} Meas(E^p)} \sum_{E^f \in \mathfrak{J}^{E^p}} \alpha_{E^f} \delta^{(m)} p_h Meas(E^f) \qquad (\forall\ E^p \in \mathfrak{T}_H^p)$$

$$\mathscr{P}\delta^{(m)}\bar{\epsilon}_H = \frac{K_{b_{E^p}}}{K_{b_{E^f}}} \frac{1}{Meas(E^p)} \int\limits_{E^p} \delta^{(m)}\bar{\epsilon}_H \qquad (\forall\ E^f \in \mathfrak{J}^{E^p}\ \forall\ E^p \in \mathscr{T}_H^p)$$

4.5.1 Convergence criterion

The convergence criterion is set in accordance with the statement of contraction (4.31), which basically states that $\sum\limits_{E^f \in \mathscr{T}_h^f} \frac{\|\delta^{(m)}\sigma_v\|_{E^f}^2}{K_{b_{E^f}}}$ keeps monotonically decreasing in successive coupling iterations. In lieu of that, the convergence criterion is to stop the iterative process after $(m+1)$ coupling iterations when

$$\frac{\sum\limits_{E^f \in \mathscr{T}_h^f} \frac{\|\delta^{(m)}\sigma_v\|_{E^f}^2}{K_{b_{E^f}}}}{\sum\limits_{E^f \in \mathscr{T}_h^f} \frac{\|\sigma_v^{(m+1)}\|_{E^f}^2}{K_{b_{E^f}}}} < TOL$$

where $TOL > 0$ is a tolerance.

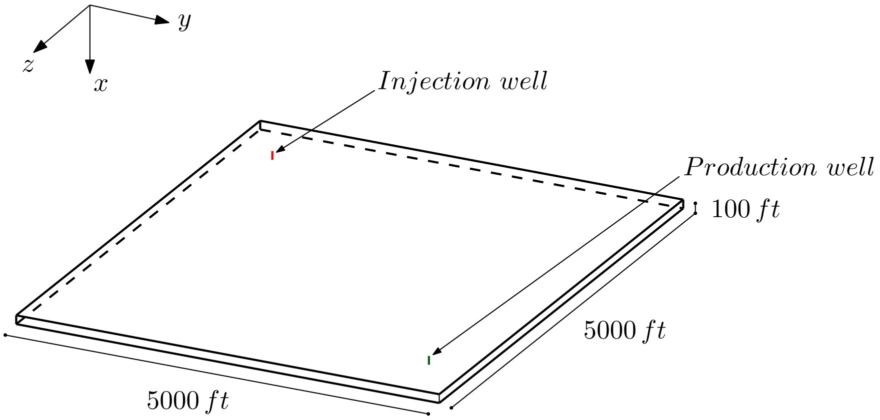

Figure 4.2: Injection well at $(500\,ft, 500\,ft)$ and production well at $(4500\,ft, 4500\,ft)$

4.6 Numerical result

The two-grid fixed stress split scheme is implemented in the in-house parallel reservoir simulator IPARS (Integrated Parallel Accurate Reservoir Simulator) at the Center for Subsurface Modeling. The bevo3 supercomputer at the Institute for Computational Engineering and Sciences is employed for the simulation.

4.6.1 Parameters

The problem domain is $\Omega \equiv 100\,ft \times 5000\,ft \times 5000\,ft$, the fine scale flow grid is $\mathscr{T}_h^f \equiv 20 \times 100 \times 100$, the coarse scale poromechanical grid is $\mathscr{T}_H^p \equiv 5 \times 25 \times 25$, fluid compressibility is $c = 3.03 \times 10^{-10}\,Pa^{-1}$, fluid viscosity is $\mu = 1.0\,cp$, reference fluid density is $\rho_0 = 62.4\,lbm/ft^3$. No flux boundary conditions are imposed for the flow subproblem. The boundary conditions for the poromechanical subproblem are zero traction on the top surface (boundary with outward normal pointing in the $-x$ direction) and zero normal displacements on the remaining boundaries. The in situ pore pressure across all flow elements is $500\,psi$. The initial porosity field is assumed homogeneous at $\phi_0 = 0.2$. The bulk modulus of the solid grains is taken to be $K_s = 15\,GPa$. The drained Young's modulus of the porous rock varies from a minimum of $0.009\,GPa$ to a maximum of $9\,GPa$ across flow elements whereas the drained poisson's ratio of the porous rock is taken to be $\nu = 0.2$ across all flow elements. As a result, the drained bulk modulus of the porous rock varies from a minimum of $\frac{0.009}{3(1-2\times 0.2)} = 0.005\,GPa$ to a maximum of $\frac{9}{3(1-2\times 0.2)} = 5\,GPa$ across flow elements and the Biot constant varies from a maximum of $1 - \frac{0.005}{15} = 0.9997$ to a minimum of $1 - \frac{5}{15} = 0.6667$ across flow elements. An injection well with constant bottom hole pressure of $1000\,psi$ is located at $(500\,ft, 500\,ft)$ and a production well with constant bottom hole pressure of $100\,psi$ is located at $(4500\,ft, 4500\,ft)$. We set the pre-specified tolerance at $TOL = 1 \times 10^{-6}$. The time step is $1\,day$ and the total simulation time is $200\,days$.

4.6.2 Theoretical versus numerically computed contraction constant

The theoretical contraction estimate γ given in Theorem 5.4.1 is computed as

$$\begin{aligned}\gamma &= \max_{E^f\in\mathscr{T}_h^f}\left(\frac{\alpha^2_{E^f}}{\frac{K_{b_{E^f}}}{M_{E^f}} + \alpha^2_{E^f}}\right) = \max_{E^f\in\mathscr{T}_h^f}\left(\frac{\alpha^2_{E^f}}{K_{b_{E^f}}\left(\phi_0 c + \frac{(\alpha_{E^f}-\phi_0)(1-\alpha_{E^f})}{K_{b_{E^f}}}\right) + \alpha^2_{E^f}}\right)\\ &= \max_{E^f\in\mathscr{T}_h^f}\left(\frac{\alpha^2_{E^f}}{K_{b_{E^f}}\phi_0 c + \alpha_{E^f} - \phi_0 + \phi_0\alpha_{E^f}}\right) = 0.9994 \qquad (4.60)\end{aligned}$$

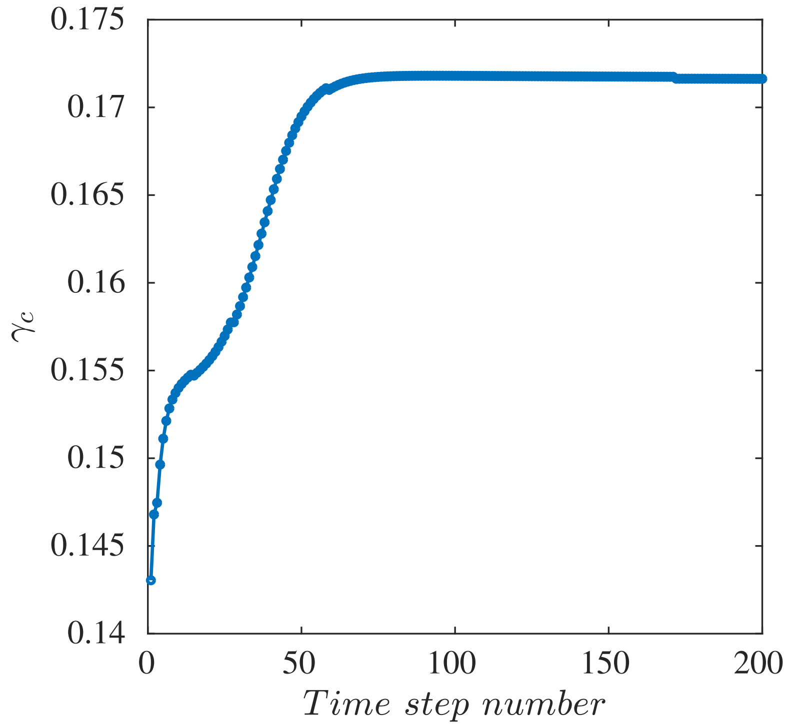

Figure 4.3: Numerically computed contraction constant

The numerically computed contraction constant γ_c for eact time step is taken to be

$$\gamma_c \equiv \max_{m=1,\cdots,m_c-1} \left[\sum_{E^f \in \mathscr{T}_h^f} \frac{\|\delta^{(m)}\sigma_v\|^2_{E^f}}{K_{b_{E^f}}} \Big/ \sum_{E^f \in \mathscr{T}_h^f} \frac{\|\delta^{(m-1)}\sigma_v\|^2_{E^f}}{K_{b_{E^f}}} \right]$$

where m_c is used to represent the number of coupling iterations to convergence at each time step. As shown in Figure 4.3, the numerically computed contraction constant gradually increases in the initial time steps and eventually settles to a value of approximately $\gamma_c = 0.172$. This result clearly indicated that the theoretical contraction estimate γ only poses an upper bound on the numerically computed contraction constant γ_c. Further, as shown in Figure **??**, the

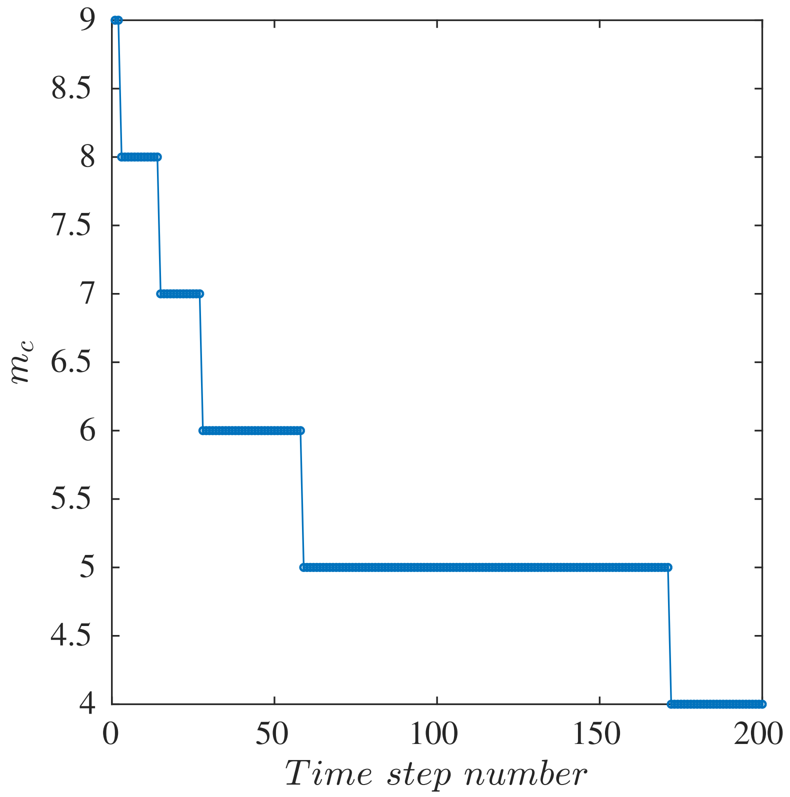

Figure 4.4: Number of coupling iterations to convergence

number of coupling iterations to convergence m_c decreases from a value of 9 to a value of 4 as we march forward in time.

4.7 Generalization of the decoupling constraint and the link with computational homogenization

In the two-grid fixed stress split scheme, the measure of mean stress that remains fixed during the flow solve is hydrostatic part of the total stress, also refered to as the mean stress. The interesting result of the work is that the convergence analysis lends itself to an expression for coarse scale bulk moduli in terms of fine scale bulk moduli, and further the coarse scale moduli are a harmonic mean of the fine scale moduli. The harmonic mean is exactly the Reuss bound (see Saeb et al. [74]) in the context of computational homogenization of multiphase composites. This observation leads to a hypothesis that there must be a measure of mean stress which when fixed during the flow solve in the two-grid approach, leads to the arithmetic mean (Voigt bound) for coarse scale bulk moduli in terms of fine scale bulk moduli. With that in mind, we define a measure of mean stress which equates to the actual mean stress only as a special case. As a result, the staggering in this work is a generalization of the fixed stress split staggering that was studied in Mikelić and Wheeler [58], Almani et al. [3] and in the above rendition.

4.7.1 The decoupling assumption

The basic idea of the two-grid staggered solution strategy is to solve the flow system (4.1)-(4.5) on a fine grid for the pressures at the current coupling iteration based on the value of a certain measure of mean stress from the previous coupling iteration. We refer to that measure of mean stress as $\bar{\sigma}$, and is expressed as follows

$$\bar{\sigma} = \eta\bar{\epsilon} - \alpha p$$

where η is an adjustable parameter, which when equated to the drained bulk modulus, lends itself to the total mean stress (refered to as σ_v) as follows

$$\bar{\sigma} = K_b\bar{\epsilon} - \alpha p \equiv \sigma_v \qquad \text{(when } \eta = K_b\text{)}$$

These pressures are then fed to the poromechanics system (2.6)-(2.11) which is solved for displacements on a coarse grid thereby updating the stress state. This updated stress state is then fed back to the flow system for the next coupling iteration. Since this strategy condemns the porous solid to follow a certain stress path during the flow solve, the convergence of the solution algorithm is not automatically guaranteed. It is important to note that the adjustable η allows for flexibility in the choice of decoupling constraint, and the fixed stress split strategy is only a special case when the adjustable parameter is identical to the drained bulk modulus i.e. when $\eta = K_b$.

4.7.2 Statement of contraction of the two-grid staggered solution algorithm

In lieu of the notations already set forth in this chapter, the measure of the mean stress is defined on the fine and coarse grids as

$$\bar{\sigma} = \eta_{E^p}\bar{\epsilon}_H - \alpha_{E^p}\mathscr{R}p_h \qquad (\forall\ E^p \in \mathscr{T}_H^p) \tag{4.61}$$

$$\bar{\sigma} = \eta_{E^f}\mathscr{P}\bar{\epsilon}_H - \alpha_{E^f}p_h \qquad (\forall\ E^f \in \mathscr{T}_h^f) \tag{4.62}$$

Theorem 4.7.1. *In the presence of medium heterogeneities, the two-grid staggered solution algorithm in which the flow subproblem is resolved on a finer grid is a contraction map with contraction constant γ and given by*

$$\sum_{E^f \in \mathscr{T}_h^f} \frac{\|\delta^{(m)}\bar{\sigma}\|^2_{E^f}}{\eta_{E^f}} + \overbrace{\sum_{E^p \in \mathscr{T}_H^p} 4G_{E^p}\|\boldsymbol{e}(\delta^{(m)}\mathbf{u}_H)\|^2_{E^p}}^{>0} + \overbrace{\sum_{E^p \in \mathscr{T}_H^p} (2K_{b_{E^p}} - \eta_{E^p})\|\delta^{(m)}\bar{\epsilon}_H\|^2_{E^p}}^{>0}$$

$$+\overbrace{\sum_{E^f\in\mathscr{T}_h^f} 2\Delta t\|\boldsymbol{\kappa}_{E^f}^{-1/2}\delta^{(m)}\mathbf{z}_h\|_{E^f}^2}^{>0} \leq \overbrace{\max_{E^f\in\mathscr{T}_h^f}\left(\frac{\alpha_{E^f}^2}{\frac{\eta_{E^f}}{M_{E^f}}+\alpha_{E^f}^2}\right)}^{\gamma<1}\sum_{E^f\in\mathscr{T}_h^f}\frac{\|\delta^{(m-1)}\bar{\sigma}\|_{E^f}^2}{\eta_{E^f}} \tag{4.63}$$

if the following conditions are satisfied

1. *First condition*

$$\sum_{E^f\in\mathscr{T}_h^f}\alpha_{E^f}(\mathscr{P}\delta^{(m)}\bar{\epsilon}_H,\delta^{(m)}p_h)_{E^f}-\sum_{E^p\in\mathscr{T}_H^p}\alpha_{E^p}(\delta^{(m)}\bar{\epsilon}_H,\mathscr{R}\delta^{(m)}p_h)_{E^p}=0 \tag{4.64}$$

2. *Second condition*

$$\sum_{E^p\in\mathscr{T}_H^p}\eta_{E^p}\|\delta^{(m)}\bar{\epsilon}_H\|_{E^p}^2-\sum_{E^f\in\mathscr{T}_h^f}\eta_{E^f}\|\mathscr{P}\delta^{(m)}\bar{\epsilon}_H\|_{E^f}^2\geq 0 \tag{4.65}$$

3. *Third condition*

$$\eta_{E^p}\leq 2K_{b_{E^p}} \qquad (\forall E^p\in\mathscr{T}_H^p) \tag{4.66}$$

Proof. • **Step 1: Flow equations**

Testing (4.26) with $\theta_h\in W_h$ such that $\theta_h|_{E^f}=\delta^{(m)}p_h\ \forall\ E^f\in\mathscr{T}_h^f$, we get

$$\begin{aligned}&\sum_{E^f\in\mathscr{T}_h^f}\varphi_{E^f}\|\delta^{(m)}p_h\|_{E^f}^2+\sum_{E^f\in\mathscr{T}_h^f}\Delta t(\nabla\cdot\delta^{(m)}\mathbf{z}_h,\delta^{(m)}p_h)_{E^f}\\&=-\sum_{E^f\in\mathscr{T}_h^f}\frac{\alpha_{E^f}}{\eta_{E^f}}(\delta^{(m-1)}\bar{\sigma},\delta^{(m)}p_h)_{E^f}\end{aligned} \tag{4.67}$$

Testing (4.27) with $\mathbf{v}_h\in\mathbf{V}_h$ such that $\mathbf{v}_h|_{E^f}\equiv\delta^{(m)}\mathbf{z}_h\ \forall\ E^f\in\mathscr{T}_h^f$, we get

$$\sum_{E^f\in\mathscr{T}_h^f}\|\boldsymbol{\kappa}_{E^f}^{-1/2}\delta^{(m)}\mathbf{z}_h\|_{E^f}^2=\sum_{E^f\in\mathscr{T}_h^f}(\delta^{(m)}p_h,\nabla\cdot\delta^{(m)}\mathbf{z}_h)_{E^f} \tag{4.68}$$

From (4.67) and (4.68), we get

$$\begin{aligned}&\sum_{E^f\in\mathscr{T}_h^f}\varphi_{E^f}\|\delta^{(m)}p_h\|^2+\sum_{E^f\in\mathscr{T}_h^f}\Delta t\|\boldsymbol{\kappa}_{E^f}^{-1/2}\delta^{(m)}\mathbf{z}_h\|_{E^f}^2\\&=-\sum_{E^f\in\mathscr{T}_h^f}\frac{\alpha_{E^f}}{\eta_{E^f}}(\delta^{(m-1)}\bar{\sigma},\delta^{(m)}p_h)_{E^f}\end{aligned}\tag{4.69}$$

• Step 2: Invoking the Young's inequality

Since the terms on the LHS of (4.69) are strictly positive, the RHS is also strictly positive. We invoke the Young's inequality

$$|ab|\le\frac{a^2}{2\varepsilon}+\frac{\varepsilon b^2}{2}\qquad\forall\ a,b,\varepsilon\in\mathbb{R},\varepsilon>0$$

for the RHS of (4.69) as follows

$$\begin{aligned}&-\frac{\alpha_{E^f}}{\eta_{E^f}}(\delta^{(m-1)}\bar{\sigma},\delta^{(m)}p_h)_{E^f}\\&\le\frac{1}{2\varepsilon_{E^f}\eta_{E^f}^2}\|\delta^{(m-1)}\bar{\sigma}\|_{E^f}^2+\frac{\varepsilon_{E^f}}{2}\|\alpha_{E^f}\delta^{(m)}p_h\|_{E^f}^2\ (\forall\ E^f\in\mathscr{T}_h^f)\end{aligned}$$

Since the above inequality is true for any $\varepsilon_{E^f}>0$, we choose $\varepsilon_{E^f}=\frac{1}{\alpha_{E^f}^2}\varphi_{E^f}$ to get

$$\begin{aligned}&-\frac{\alpha_{E^f}}{\eta_{E^f}}(\delta^{(m-1)}\bar{\sigma},\delta^{(m)}p_h)_{E^f}\\&\le\frac{\alpha_{E^f}^2}{2\eta_{E^f}\varphi_{E^f}}\frac{\|\delta^{(m-1)}\bar{\sigma}\|_{E^f}^2}{\eta_{E^f}}+\frac{\varphi_{E^f}}{2}\|\delta^{(m)}p_h\|_{E^f}^2\ (\forall\ E^f\in\mathscr{T}_h^f)\end{aligned}$$

In lieu of the above, (4.69) is written as

$$\begin{aligned}&\sum_{E^f\in\mathscr{T}_h^f}\varphi_{E^f}\|\delta^{(m)}p_h\|_{E^f}^2+\sum_{E^f\in\mathscr{T}_h^f}\Delta t\|\boldsymbol{\kappa}_{E^f}^{-1/2}\delta^{(m)}\mathbf{z}_h\|_{E^f}^2\\&\le\sum_{E^f\in\mathscr{T}_h^f}\frac{\alpha_{E^f}^2}{2\eta_{E^f}\varphi_{E^f}}\frac{\|\delta^{(m-1)}\bar{\sigma}\|_{E^f}^2}{\eta_{E^f}}+\sum_{E^f\in\mathscr{T}_h^f}\frac{\varphi_{E^f}}{2}\|\delta^{(m)}p_h\|_{E^f}^2\end{aligned}$$

which can also be written as

$$\sum_{E^f \in \mathscr{T}_h^f} \frac{\varphi_{E^f}}{2} \|\delta^{(m)} p_h\|^2 + \sum_{E^f \in \mathscr{T}_h^f} \Delta t \|\boldsymbol{\kappa}_{E^f}^{-1/2} \delta^{(m)} \mathbf{z}_h\|_{E^f}^2$$
$$\leq \sum_{E^f \in \mathscr{T}_h^f} \frac{\alpha_{E^f}^2}{2\eta_{E^f} \varphi_{E^f}} \frac{\|\delta^{(m-1)} \bar{\sigma}\|_{E^f}^2}{\eta_{E^f}}$$

which, after noting that $\varphi_{E^f} \equiv \left(\frac{1}{M_{E^f}} + \frac{\alpha_{E^f}^2}{\eta_{E^f}} \right) > \frac{\alpha_{E^f}^2}{\eta_{E^f}}$, can also be written as

$$\sum_{E^f \in \mathscr{T}_h^f} \frac{\alpha_{E^f}^2}{\eta_{E^f}} \|\delta^{(m)} p_h\|_{E^f}^2 + \sum_{E^f \in \mathscr{T}_h^f} 2\Delta t \|\boldsymbol{\kappa}_{E^f}^{-1/2} \delta^{(m)} \mathbf{z}_h\|_{E^f}^2$$
$$\leq \sum_{E^f \in \mathscr{T}_h^f} \frac{\alpha_{E^f}^2}{\eta_{E^f} \varphi_{E^f}} \frac{\|\delta^{(m-1)} \bar{\sigma}\|_{E^f}^2}{\eta_{E^f}} \tag{4.70}$$

- **Step 3: Poromechanics equations**

Testing (4.28) with $\mathbf{q}_H \in \mathbf{Q}_H$ such that $\mathbf{q}|_{E^p} = 2\delta^{(m)} \mathbf{u}_H \ \forall \ E^p \in \mathscr{T}_H^p$ and noting that $\nabla \cdot \delta^{(m)} \mathbf{u}_H \equiv \delta^{(m)} \bar{\epsilon}_H$, we get

$$\sum_{E^p \in \mathscr{T}_H^p} 4G_{E^p} \|\boldsymbol{e}(\delta^{(m)} \mathbf{u}_H)\|_{E^p}^2 + \sum_{E^p \in \mathscr{T}_H^p} 2(\delta^{(m)} \sigma_v, \delta^{(m)} \bar{\epsilon}_H)_{E^p} = 0 \tag{4.71}$$

Further, from (4.61), we note that $\delta^{(m)} \bar{\sigma} = K_{b_{E^p}} \delta^{(m)} \bar{\epsilon}_H - \alpha_{E^p} \mathscr{R} \delta^{(m)} p_h \ \forall \ E^p \in \mathscr{T}_H^p$. As a result, (4.71) is written as

$$\sum_{E^p \in \mathscr{T}_H^p} 4G_{E^p} \|\boldsymbol{e}(\delta^{(m)} \mathbf{u}_H)\|_{E^p}^2 + \sum_{E^p \in \mathscr{T}_H^p} 2K_{b_{E^p}} \|\delta^{(m)} \bar{\epsilon}_H\|_{E^p}^2$$
$$- \sum_{E^p \in \mathscr{T}_H^p} 2\alpha_{E^p} (\delta^{(m)} \bar{\epsilon}_H, \mathscr{R} \delta^{(m)} p_h)_{E^p} = 0 \tag{4.72}$$

- **Step 4: Combining flow and poromechanics equations**

Adding (4.70) and (4.72), we get

$$\sum_{E^f \in \mathscr{T}_h^f} \frac{\alpha_{E^f}^2}{\eta_{E^f}} \|\delta^{(m)} p_h\|_{E^f}^2 + \sum_{E^f \in \mathscr{T}_h^f} 2\Delta t \|\boldsymbol{\kappa}_{E^f}^{-1/2} \delta^{(m)} \mathbf{z}_h\|_{E^f}^2$$

$$
\begin{aligned}
&+ \sum_{E^p \in \mathscr{T}_H^p} 4G_{E^p} \|\boldsymbol{e}(\delta^{(m)}\mathbf{u}_H)\|_{E^p}^2 + \sum_{E^p \in \mathscr{T}_H^p} 2K_{b_{E^p}} \|\delta^{(m)}\bar{\epsilon}_H\|_{E^p}^2 \\
&- \sum_{E^p \in \mathscr{T}_H^p} 2\alpha_{E^p}(\delta^{(m)}\bar{\epsilon}_H, \mathscr{R}\delta^{(m)}p_h)_{E^p} \leq \sum_{E^f \in \mathscr{T}_h^f} \frac{\alpha_{E^f}^2}{\eta_{E^f}\varphi_{E^f}} \frac{\|\delta^{(m-1)}\bar{\sigma}\|_{E^f}^2}{\eta_{E^f}} \qquad (4.73)
\end{aligned}
$$

Now, from (4.62), we note that

$$
\|\delta^{(m)}\bar{\sigma}\|_{E^f}^2 = \alpha_{E^f}^2 \|\delta^{(m)}p_h\|_{E^f}^2 + \eta_{E^f}^2 \|\mathscr{P}\delta^{(m)}\bar{\epsilon}_H\|_{E^f}^2 - 2\eta_{E^f}\alpha_{E^f}(\mathscr{P}\delta^{(m)}\bar{\epsilon}_H, \delta^{(m)}p_h)_{E^f}
$$

$$
(\forall\ E^f \in \mathscr{T}_h^f)
$$

which implies that

$$
\begin{aligned}
&\frac{\alpha_{E^f}^2}{\eta_{E^f}} \|\delta^{(m)}p_h\|_{E^f}^2 = \frac{\|\delta^{(m)}\bar{\sigma}\|_{E^f}^2}{\eta_{E^f}} - \eta_{E^f} \|\mathscr{P}\delta^{(m)}\bar{\epsilon}_H\|_{E^f}^2 + 2\alpha_{E^f}(\mathscr{P}\delta^{(m)}\bar{\epsilon}_H, \delta^{(m)}p_h)_{E^f} \\
&(\forall\ E^f \in \mathscr{T}_h^f) \qquad (4.74)
\end{aligned}
$$

Substituting (4.74) in (4.73), we get

$$
\begin{aligned}
&\sum_{E^f \in \mathscr{T}_h^f} \frac{\|\delta^{(m)}\bar{\sigma}\|_{E^f}^2}{\eta_{E^f}} + \overbrace{\sum_{E^p \in \mathscr{T}_H^p} 4G_{E^p} \|\boldsymbol{e}(\delta^{(m)}\mathbf{u}_H)\|_{E^p}^2}^{>0} + \overbrace{\sum_{E^p \in \mathscr{T}_H^p} (2K_{b_{E^p}} - \eta_{E^p}) \|\delta^{(m)}\bar{\epsilon}_H\|_{E^p}^2}^{>0} \\
&+ \left[\overbrace{\sum_{E^f \in \mathscr{T}_h^f} 2\alpha_{E^f}(\mathscr{P}\delta^{(m)}\bar{\epsilon}_H, \delta^{(m)}p_h)_{E^f} - \sum_{E^p \in \mathscr{T}_H^p} 2\alpha_{E^p}(\delta^{(m)}\bar{\epsilon}_H, \mathscr{R}\delta^{(m)}p_h)_{E^p}}^{\text{Set } = 0 \text{ to obtain expressions for } \eta_{\mathrm{E^p}}\ \forall\ \mathrm{E^p} \in \mathscr{T}_{\mathrm{H}}^{\mathrm{p}} \text{ and } \mathscr{P}\delta^{(\mathrm{m})}\bar{\epsilon}_{\mathrm{H}}\ \forall\ \mathrm{E^f} \in \mathscr{T}_{\mathrm{h}}^{\mathrm{f}}} \right] \\
&+ \left[\overbrace{\sum_{E^p \in \mathscr{T}_H^p} \eta_{E^p} \|\delta^{(m)}\bar{\epsilon}_H\|_{E^p}^2 - \sum_{E^f \in \mathscr{T}_h^f} \eta_{E^f} \|\mathscr{P}\delta^{(m)}\bar{\epsilon}_H\|_{E^f}^2}^{\text{Turns out to be } \geq 0 \text{ in lieu of Cauchy}-\text{Schwartz inequality}} \right] \\
&+ \overbrace{\sum_{E^f \in \mathscr{T}_h^f} 2\Delta t \|\boldsymbol{\kappa}_{E^f}^{-1/2}\delta^{(m)}\mathbf{z}_h\|_{E^f}^2}^{>0} \leq \gamma \sum_{E^f \in \mathscr{T}_h^f} \frac{\|\delta^{(m-1)}\bar{\sigma}\|_{E^f}^2}{\eta_{E^f}} \qquad (4.75)
\end{aligned}
$$

The statement (4.75) is a contraction map in a sense that

$$
\sum_{E^f \in \mathscr{T}_h^f} \frac{\|\delta^{(0)}\bar{\sigma}\|_{E^f}^2}{\eta_{E^f}} > \sum_{E^f \in \mathscr{T}_h^f} \frac{\|\delta^{(1)}\bar{\sigma}\|_{E^f}^2}{\eta_{E^f}} > \sum_{E^f \in \mathscr{T}_h^f} \frac{\|\delta^{(2)}\bar{\sigma}\|_{E^f}^2}{\eta_{E^f}} > ...
$$

with contraction constant γ given by

$$\gamma \equiv \max_{E^f \in \mathscr{T}_h^f} \left(\frac{\alpha_{E^f}^2}{\eta_{E^f} \varphi_{E^f}} \right) = \max_{E^f \in \mathscr{T}_h^f} \left(\frac{\alpha_{E^f}^2}{\frac{\eta_{E^f}}{M_{E^f}} + \alpha_{E^f}^2} \right) < 1$$

provided the following are true

$$\sum_{E^f \in \mathscr{T}_h^f} \alpha_{E^f} (\mathscr{P}\delta^{(m)}\bar{\epsilon}_H, \delta^{(m)} p_h)_{E^f} - \sum_{E^p \in \mathscr{T}_H^p} \alpha_{E^p} (\delta^{(m)}\bar{\epsilon}_H, \mathscr{R}\delta^{(m)} p_h)_{E^p} = 0 \tag{4.76}$$

$$\sum_{E^p \in \mathscr{T}_H^p} \eta_{E^p} \|\delta^{(m)}\bar{\epsilon}_H\|_{E^p}^2 - \sum_{E^f \in \mathscr{T}_h^f} \eta_{E^f} \|\mathscr{P}\delta^{(m)}\bar{\epsilon}_H\|_{E^f}^2 \geq 0 \tag{4.77}$$

$$\eta_{E^p} \leq 2K_{b_{E^p}} \qquad (\forall E^p \in \mathscr{T}_H^p) \tag{4.78}$$

□

Corollary 4.7.1.1. *Satisfaction of the decoupling constraint during the flow solve at both scales leads to the following expressions for the upscaled pore pressures*

$$\mathscr{R}\delta^{(m)} p_h = \frac{\eta_{E^p}}{\alpha_{E^p}} \sum_{E^f \in \mathfrak{J}^{E^p}} \frac{\alpha_{E^f}}{\eta_{E^f}} \delta^{(m)} p_h \frac{Meas(E^f)}{Meas(E^p)} \qquad (\forall\ E^p \in \mathfrak{T}_H^p)$$

Proof. • **Step 1: Using the fact that pore pressure is frozen during the poromechanical solve**

Since the pore pressure is frozen during the poromechanical solve, the total pore pressure change in a coupling iteration is the same as the pore pressure change calculated during the flow solve in the coupling iteration as follows

$$\mathscr{R}\delta_f^{(m)} p_h = \mathscr{R}\delta^{(m)} p_h \qquad (\forall\ E^p \in \mathscr{T}_H^p) \tag{4.79}$$

$$\delta_f^{(m)} p_h = \delta^{(m)} p_h \qquad (\forall\ E^f \in \mathscr{T}_h^f) \tag{4.80}$$

- **Step 2: Applying the decoupling constraint on both scales**

Now, the decoupling constraint implies that there is no change in the measure of the mean stress of the system during the flow solve. This naturally implies that

$$\int\limits_{E^p} \delta_f^{(m)} \bar{\sigma} = 0 \qquad (\forall\ E^p \in \mathscr{T}_H^p)$$

In lieu of (4.61), we write the above as

$$\int\limits_{E^p} (\eta_{E^p} \delta_f^{(m)} \bar{\epsilon}_H - \alpha_{E^p} \mathscr{R} \delta_f^{(m)} p_h) = 0 \qquad (\forall\ E^p \in \mathscr{T}_H^p)$$

which, in lieu of (4.79), can be written as

$$\int\limits_{E^p} \delta_f^{(m)} \bar{\epsilon}_H = \frac{\alpha_{E^p}}{\eta_{E^p}} \mathscr{R} \delta^{(m)} p_h Meas(E^p) \qquad (\forall\ E^p \in \mathscr{T}_H^p) \tag{4.81}$$

Denoting $\int\limits_{E^f} \delta_f^{(m)} \mathscr{P} \bar{\epsilon}_H$ is the change in volume of each element E^f of $\mathcal{J}^{E^p}$, we now impose the decoupling constraint on each element E^f of $\mathcal{J}^{E^p}$ as follows

$$\int\limits_{E^f} \delta_f^{(m)} \bar{\sigma} \equiv \int\limits_{E^f} (\eta_{E^f} \mathscr{P} \delta_f^{(m)} \bar{\epsilon}_H - \alpha_{E^f} \delta^{(m)} p_h) = 0 \qquad (\forall\ E^f \in \mathcal{J}^{E^p})$$

which, in lieu of (4.80), can be written as

$$\int\limits_{E^f} \mathscr{P} \delta_f^{(m)} \bar{\epsilon}_H = \frac{\alpha_{E^f}}{\eta_{E^f}} \delta^{(m)} p_h Meas(E^f) \qquad (\forall\ E^f \in \mathscr{T}_h^f) \tag{4.82}$$

- **Step 3: Using the fact that the change in volume measured on both scales should be identical**

The term $\int\limits_{E^p} \delta_f^{(m)} \bar{\epsilon}_H$ is the change in volume of E^p during the flow solve in the

$(m+1)^{th}$ coupling iteration. This naturally equates the sum of corresponding changes in volumes of the elements of $\mathfrak{I}^{E^p}$ as follows

$$\int_{E^p} \delta_f^{(m)} \bar{\epsilon}_H \equiv \sum_{E^f \in \mathfrak{I}^{E^p}} \int_{E^f} \mathscr{P} \delta_f^{(m)} \bar{\epsilon}_H \qquad (\forall\ E^p \in \mathscr{T}_H^p) \tag{4.83}$$

From (4.82) and (4.83), we get

$$\int_{E^p} \delta_f^{(m)} \bar{\epsilon}_H = \sum_{E^f \in \mathfrak{I}^{E^p}} \frac{\alpha_{E^f}}{\eta_{E^f}} \delta^{(m)} p_h Meas(E^f) \qquad (\forall\ E^p \in \mathscr{T}_H^p) \tag{4.84}$$

From (4.81) and (4.84), we get

$$\sum_{E^f \in \mathfrak{I}^{E^p}} \frac{\alpha_{E^f}}{\eta_{E^f}} \delta^{(m)} p_h Meas(E^f) = \frac{\alpha_{E^p}}{\eta_{E^p}} \mathscr{R} \delta^{(m)} p_h Meas(E^p) \qquad (\forall\ E^p \in \mathscr{T}_H^p)$$

which results in

$$\mathscr{R} \delta^{(m)} p_h = \frac{\eta_{E^p}}{\alpha_{E^p}} \sum_{E^f \in \mathfrak{I}^{E^p}} \frac{\alpha_{E^f}}{\eta_{E^f}} \delta^{(m)} p_h \frac{Meas(E^f)}{Meas(E^p)} \qquad (\forall\ E^p \in \mathscr{T}_H^p) \tag{4.85}$$

□

Corollary 4.7.1.2. *Satisfaction of the condition* (4.64) *leads to the following expressions for the effective bulk moduli for the coarse scale poromechanical solve*

$$\eta_{E^p} = \frac{1}{\sum\limits_{E^f \in \mathfrak{I}^{E^p}} \frac{1}{\eta_{E^f}} \frac{Meas(E^f)}{Meas(E^p)}} \qquad (\forall\ E^p \in \mathscr{T}_H^p)$$

and the following expressions for the downscaled volumetric strains

$$\mathscr{P} \delta^{(m)} \bar{\epsilon}_H = \frac{\eta_{E^p}}{\eta_{E^f}} \frac{1}{Meas(E^p)} \int_{E^p} \delta^{(m)} \bar{\epsilon}_H \qquad (\forall\ E^f \in \mathfrak{I}^{E^p}\ \forall\ E^p \in \mathscr{T}_H^p)$$

Proof. • **Step 1: Recasting the first term on LHS of** (4.64)

We start by modifying the first term on LHS of (4.64) as follows

$$\sum_{E^f \in \mathscr{T}_h^f} \alpha_{E^f} (\mathscr{P}\delta^{(m)}\bar{\epsilon}_H, \delta^{(m)}p_h)_{E^f} = \sum_{E^f \in \mathscr{T}_h^f} \alpha_{E^f} \delta^{(m)}p_h \mathscr{P}\delta^{(m)}\bar{\epsilon}_H Meas(E^f) \tag{4.86}$$

where we note that $\delta^{(m)}p_h \in W_h$. Since a flow element $E^f \in \mathscr{T}_h^f$ in uniquely associated with a poromechanical element E^p via $\mathfrak{I}^{E^p}$, we can write

$$\sum_{E^f \in \mathscr{T}_h^f} \alpha_{E^f} \delta^{(m)}p_h \mathscr{P}\delta^{(m)}\bar{\epsilon}_H Meas(E^f) = \sum_{E^p \in \mathscr{T}_H^p} \sum_{E^f \in \mathfrak{I}^{E^p}} \alpha_{E^f} \delta^{(m)}p_h \mathscr{P}\delta^{(m)}\bar{\epsilon}_H Meas(E^f)$$

In lieu of the above, we write (4.86) as

$$\sum_{E^f \in \mathscr{T}_h^f} \alpha_{E^f} (\mathscr{P}\delta^{(m)}\bar{\epsilon}_H, \delta^{(m)}p_h)_{E^f} = \sum_{E^p \in \mathscr{T}_H^p} \sum_{E^f \in \mathfrak{I}^{E^p}} \alpha_{E^f} \delta^{(m)}p_h \mathscr{P}\delta^{(m)}\bar{\epsilon}_H Meas(E^f) \tag{4.87}$$

• **Step 2: Recasting the second term on LHS of** (4.64)

Next, we modify the second term on LHS of (4.64) as follows

$$\sum_{E^p \in \mathscr{T}_H^p} \alpha_{E^p} (\delta^{(m)}\bar{\epsilon}_H, \mathscr{R}\delta^{(m)}p_h)_{E^p} \equiv \sum_{E^p \in \mathscr{T}_H^p} 2\alpha_{E^p} \mathscr{R}\delta^{(m)}p_h \int_{E^p} \delta^{(m)}\bar{\epsilon}_H \tag{4.88}$$

where we note that $\mathscr{R}\delta^{(m)}p_h \in W_H$. In lieu of (4.87) and (4.88), the first condition given by (4.64) is rewritten as

$$\sum_{E^p \in \mathscr{T}_H^p} \sum_{E^f \in \mathfrak{I}^{E^p}} \alpha_{E^f} \delta^{(m)}p_h \mathscr{P}\delta^{(m)}\bar{\epsilon}_H Meas(E^f) = \sum_{E^p \in \mathscr{T}_H^p} \alpha_{E^p} \mathscr{R}\delta^{(m)}p_h \int_{E^p} \delta^{(m)}\bar{\epsilon}_H \tag{4.89}$$

• **Step 3: Substituting the expression for upscaled pore pressures**

Substituting the expression (4.85) for the upscaled pore pressure in (4.89), we get

$$\sum_{E^p \in \mathscr{T}_H^p} \sum_{E^f \in \mathfrak{I}^{E^p}} \alpha_{E^f} \delta^{(m)} p_h \mathscr{P} \delta^{(m)} \bar{\epsilon}_H Meas(E^f)$$

$$= \sum_{E^p \in \mathscr{T}_H^p} \alpha_{E^p} \overbrace{\frac{\eta_{E^p}}{\alpha_{E^p}} \sum_{E^f \in \mathfrak{I}^{E^p}} \frac{\alpha_{E^f}}{\eta_{E^f}} \delta^{(m)} p_h \frac{Meas(E^f)}{Meas(E^p)}}^{\mathscr{R}\delta^{(m)} p_h} \int\limits_{E^p} \delta^{(m)} \bar{\epsilon}_H$$

which implies that

$$\sum_{E^p \in \mathscr{T}_H^p} \sum_{E^f \in \mathfrak{I}^{E^p}} \left(\mathscr{P} \delta^{(m)} \bar{\epsilon}_H - \frac{\eta_{E^p}}{\eta_{E^f}} \frac{1}{Meas(E^p)} \int\limits_{E^p} \delta^{(m)} \bar{\epsilon}_H \right) \alpha_{E^f} \delta^{(m)} p_h Meas(E^f) = 0$$

which, in lieu of the linear independence of the basis p_h ($\forall\ E^f \in \mathscr{T}_h^f$) of the pressure space on the fine scale flow grid, implies that

$$\mathscr{P} \delta^{(m)} \bar{\epsilon}_H - \frac{\eta_{E^p}}{\eta_{E^f}} \frac{1}{Meas(E^p)} \int\limits_{E^p} \delta^{(m)} \bar{\epsilon}_H = 0 \qquad (\forall\ E^f \in \mathfrak{I}^{E^p}\ \forall\ E^p \in \mathscr{T}_H^p)$$

implying that

$$\mathscr{P} \delta^{(m)} \bar{\epsilon}_H = \frac{\eta_{E^p}}{\eta_{E^f}} \frac{1}{Meas(E^p)} \int\limits_{E^p} \delta^{(m)} \bar{\epsilon}_H \qquad (\forall\ E^f \in \mathfrak{I}^{E^p}\ \forall\ E^p \in \mathscr{T}_H^p) \tag{4.90}$$

• **Step 4: Using the fact that the change in volume measured on both scales should be identical**

The change in volume of E^p over the $(m+1)^{th}$ coupling iteration equates the sum of corresponding changes in volumes of the elements of $\mathfrak{I}^{E^p}$ as follows

$$\int\limits_{E^p} \delta^{(m)} \bar{\epsilon}_H = \sum_{E^f \in \mathfrak{I}^{E^p}} \int\limits_{E^f} \mathscr{P} \delta^{(m)} \bar{\epsilon}_H = \sum_{E^f \in \mathfrak{I}^{E^p}} \mathscr{P} \delta^{(m)} \bar{\epsilon}_H Meas(E^f) \qquad (\forall\ E^p \in \mathscr{T}_H^p) \tag{4.91}$$

In lieu of (4.90) and (4.91), we get

$$\int\limits_{E^p} \delta^{(m)}\bar{\epsilon}_H = \int\limits_{E^p} \delta^{(m)}\bar{\epsilon}_H \sum_{E^f \in \mathfrak{I}^{E^p}} \frac{\eta_{E^p}}{\eta_{E^f}} \frac{Meas(E^f)}{Meas(E^p)} \qquad (\forall\ E^p \in \mathscr{T}_H^p)$$

which finally leads to

$$\frac{1}{\eta_{E^p}} = \sum_{E^f \in \mathfrak{I}^{E^p}} \frac{1}{\eta_{E^f}} \frac{Meas(E^f)}{Meas(E^p)} \qquad (\forall\ E^p \in \mathscr{T}_H^p) \tag{4.92}$$

□

4.7.3 The Voigt bound, the Reuss bound and the contraction constant

The contraction constant is given by

$$\gamma = \max_{E^f \in \mathscr{T}_h^f} \left(\frac{\alpha_{E^f}^2}{\frac{\eta_{E^f}}{M_{E^f}} + \alpha_{E^f}^2} \right) < 1$$

It is clear to see that the minimum value of contraction constant is obtained when the adjustable parameter takes the maximum possible value. To interrogate the maximum value that the adjustable parameter can achieve, we look at the third condition for the satisfaction of the contractivity given by

$$\eta_{E^p} \leq 2K_{b_{E^p}} \qquad (\forall E^p \in \mathscr{T}_H^p)$$

It is clear when $\eta = 2K_b$, we obtain the minimum contraction constant thus implying fastest convergence of the staggered solution algorithm. The expression (4.59) for the coarse scale moduli in terms on fine scale data is given by

$$\frac{1}{\eta_{E^p}} = \sum_{E^f \in \mathfrak{I}^{E^p}} \frac{1}{\eta_{E^f}} \frac{Meas(E^f)}{Meas(E^p)} \qquad (\forall\ E^p \in \mathscr{T}_H^p) \tag{4.93}$$

The following cases arise

- The adjustable parameter is equal to twice the drained bulk modulus i.e. $\eta \equiv 2K_b$

  $$\frac{1}{2K_{b_{E^p}}} = \sum_{E^f \in \mathfrak{I}^{E^p}} \frac{1}{2K_{b_{E^f}}} \frac{Meas(E^f)}{Meas(E^p)} \qquad (\forall\ E^p \in \mathscr{T}_H^p)$$

  $$\implies \frac{1}{K_{b_{E^p}}} = \sum_{E^f \in \mathfrak{I}^{E^p}} \frac{1}{K_{b_{E^f}}} \frac{Meas(E^f)}{Meas(E^p)} \qquad (\forall\ E^p \in \mathscr{T}_H^p)$$

  In this case, the coarse scale bulk moduli are harmonic mean of the fine scale data, thus representing the Reuss bound

- The adjustable parameter is equal to inverse of the drained bulk modulus i.e. $\eta \equiv \frac{1}{K_b}$

  $$K_{b_{E^p}} = \sum_{E^f \in \mathfrak{I}^{E^p}} K_{b_{E^f}} \frac{Meas(E^f)}{Meas(E^p)} \qquad (\forall\ E^p \in \mathscr{T}_H^p)$$

  In this case, the coarse scale bulk moduli are arithmetic mean of the fine scale data, thus representing the Voigt bound

We already know that the Reuss and Voigt bounds on effective moduli yield the lower and upper bounds for the elastic strain energy for multiphase composites respectively (see Saeb et al. [74]). In lieu of that, we state that the adjustable parameter is bounded above by the drained bulk modulus and below by the inverse of bulk modulus as follows

$$\frac{1}{K_b} \leq \eta \leq 2K_b$$

4.8 Summary

We invoke the concept of contraction mapping to establish the theoretical convergence of the solution algorithm presented in previous chapter and further generalize the algorithm in the presence of heterogeneities in the porous medium. It is important to note that the analysis holds for the linearized flow model coupled with the linear poromechanics model. We provide a module that explains how we start with the mass conservation equation for the slightly compressible single phase flow model and make linearity assumptions to arrive at the mass conservation equation for the linearized flow model. It is also important to note that the convergence analysis is performed for degenerate case of nested grids for flow and poromechanics. Later, we establish the link between the decoupling constraint for the two-grid algorithm and the concept of computational homogenization of multiphase composites.

Chapter 5

Convergence analysis of the fixed stress split algorithm for anisotropic poroelasticity with flow and poromechanics being resolved on the same grid

The subject of this chapter is the extension of the decoupling assumption to the case of anisotropic poroelasticity with tensor Biot parameter. The micromechanical analyses for the case of anisotropic poroelasticity (see Carroll [18], Carroll and Katsube [19], Katsube [52], Thompson and Willis [86]) revealed that the modification to the stress applied to the porous solid due to the presence of pore fluid pressure is not hydrostatic, as it is in the case of isotropic poroelasticity (see Nur and Byerlee [62]). Further, unlike in case of isotropic poroelasticity where the solid-fluid coupling parameter is a scalar (see Biot [8], Geertsma [36], Skempton [80], Nur and Byerlee [62]), the coupling parameter in case of anisotropic poroelasticity is a tensor (see Biot [9], Coussy [24]). Intuitively, this implies that the decoupling assumption of freezing the hydrostatic part of the stress tensor during the flow solve in every coupling iteration for the case of isotropic poroelasticity requires a generalization for the case of anisotropic poroelasticity. As we shall show in the ensuing convergence analysis, the decoupling assumption of freezing all components of

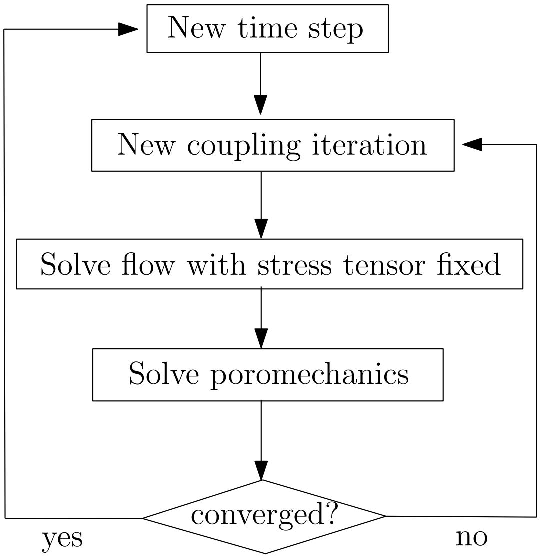

Figure 5.1: Fixed stress split iterative scheme for anisotropic poroelasticity with tensor Biot parameter

the stress tensor during the flow solve in every coupling iteration does enjoy theoretical convergence for the case of anisotropic poroelasticity with tensor Biot parameter. After that, we shall proceed to use the Mandel's problem analytical solution (see Abousleiman et al. [1]) for the transversely isotropic case to show that the decoupling assumption does enjoy numerical convergence as well.

5.0.1 Preliminaries

The inner product of two second order tensors **S** and **T** is given by (see Gurtin et al. [41])

$$\mathbf{S} : \mathbf{T} = S_{ij}T_{ij} \qquad (i, j = 1, 2, 3)$$

A fourth order tensor is a linear transformation of a second order tensor to a second order tensor in the following manner (see Gurtin et al. [41])

$$\mathbb{P}\mathbf{S} = \mathbf{T} \rightarrow \mathbb{P}_{ijkl}S_{kl} = T_{ij} \qquad (i,j,k,l = 1,2,3)$$

5.1 Flow model

The equations of the flow model are already given in Module 4.1.

5.2 Poromechanics model

The equations of the poromechanics model are already in Module 2.1.2. The generalized Hooke's law (see Coussy [24])

$$\boldsymbol{\sigma} = \mathbb{M}\boldsymbol{\epsilon} - \boldsymbol{\alpha}p \tag{5.1}$$

where $\mathbb{M}$ is the fourth order anisotropic elasticity tensor and $\boldsymbol{\alpha}$ is the Biot tensor. The inverse of the generalized Hooke's law (5.1) is given by (see Cheng [23])

$$\boldsymbol{\epsilon} = \mathbb{C}\boldsymbol{\sigma} + \frac{1}{3}C\mathbf{B}p \tag{5.2}$$

where $\mathbb{C}$ is the fourth order anisotropic compliance tensor, $C(>0)$ is a generalized Hooke's law constant and $\mathbf{B}$ is a generalization of the Skempton pore pressure coefficient B (see Skempton [79]) for anisotropic poroelasticity.

5.3 Fluid content

The fluid content ζ is given by (see Cheng [23])

$$\zeta = Cp + \frac{1}{3}C\mathbf{B} : \boldsymbol{\sigma} \equiv \frac{1}{M}p + \boldsymbol{\alpha} : \boldsymbol{\epsilon} \tag{5.3}$$

where $M(> 0)$ is a generalization of the Biot modulus (see Biot and Willis [10]) for anisotropic poroelasticity.

5.4 Statement of contraction of the fixed stress split scheme for anisotropic poroelasticity with Biot tensor

Let $\mathfrak{T}_h$ be finite element partition of Ω consisting of distorted hexahedral elements E where $h = \max_{E \in \mathfrak{T}_h} diam(E)$.

5.4.1 Discrete variational statements for the flow subproblem

Before arriving at the discrete variational statement of the flow model, we impose the fixed stress constraint on the strong form of the mass conservation equation (4.1). In lieu of (5.3), we write (4.1) as

$$\begin{gathered}\frac{\partial}{\partial t}(Cp + \frac{C}{3}\mathbf{B} : \boldsymbol{\sigma}) + \nabla \cdot \mathbf{z} = q \\ C\frac{\partial p}{\partial t} + \nabla \cdot \mathbf{z} = q - \frac{C}{3}\mathbf{B} : \frac{\partial \boldsymbol{\sigma}}{\partial t}\end{gathered} \tag{5.4}$$

Using backward Euler in time, the discrete in time form of (5.4) for the m^{th} coupling iteration in the $(n+1)^{th}$ time step is written as

$$C\frac{1}{\Delta t}(p^{m,n+1} - p^n) + \nabla \cdot \mathbf{z}^{m,n+1} = q^{n+1} - \frac{1}{\Delta t}\frac{C}{3}\mathbf{B} : (\boldsymbol{\sigma}^{m,n+1} - \boldsymbol{\sigma}^n)$$

where Δt is the time step and the source term as well as the terms evaluated at the previous time level n do not depend on the coupling iteration count as they are known quantities. The fixed stress constraint implies that $\boldsymbol{\sigma}^{m,n+1}$ gets replaced by $\boldsymbol{\sigma}^{m-1,n+1}$ i.e. the computation of $p^{m,n+1}$ and $\mathbf{z}^{m,n+1}$ is based on the value of stress updated after the poromechanics solve from the previous coupling iteration $m-1$ at the current time level $n+1$. The modified equation is written as

$$C(p^{m,n+1}-p^n)+\Delta t\nabla\cdot\mathbf{z}^{m,n+1}=\Delta t q^{n+1}-\frac{C}{3}\mathbf{B}:(\boldsymbol{\sigma}^{m,n+1}-\boldsymbol{\sigma}^n)$$

As a result, the discrete weak form of (4.1) is given by

$$\begin{aligned}&C(p_h^{m,n+1}-p_h^n,\theta_h)_\Omega+\Delta t(\nabla\cdot\mathbf{z}_h^{m,n+1},\theta_h)_\Omega\\&=\Delta t(q^{n+1},\theta_h)_\Omega-\frac{C}{3}(\mathbf{B}:(\boldsymbol{\sigma}^{m-1,n+1}-\boldsymbol{\sigma}^n),\theta_h)_\Omega\end{aligned}$$

Replacing m by $m+1$ and subtracting the two equations, we get

$$C(\delta^{(m)}p_h,\theta_h)_\Omega+\Delta t(\nabla\cdot\delta^{(m)}\mathbf{z}_h,\theta_h)_\Omega=-\frac{C}{3}(\mathbf{B}:\delta^{(m-1)}\boldsymbol{\sigma},\theta_h)_\Omega$$

The weak form of the Darcy law (4.2) for the m^{th} coupling iteration in the $(n+1)^{th}$ time step is given by

$$(\boldsymbol{\kappa}^{-1}\mathbf{z}^{m,n+1},\mathbf{v})_\Omega=-(\nabla p^{m,n+1},\mathbf{v})_\Omega+(\rho_0\mathbf{g},\mathbf{v})_\Omega\ \forall\ \mathbf{v}\in\mathbf{V}(\Omega) \tag{5.5}$$

where $\mathbf{V}(\Omega)$ is given by

$$\mathbf{V}(\Omega)\equiv\mathbf{H}(div,\Omega)\cap\left\{\mathbf{v}:\mathbf{v}\cdot\mathbf{n}=0\text{ on }\Gamma_N^f\right\}$$

We use the divergence theorem to evaluate the first term on RHS of (5.5) as follows

$$
\begin{aligned}
&(\nabla p^{m,n+1}, \mathbf{v})_\Omega = (\nabla, p^{m,n+1}\mathbf{v})_\Omega - (p^{m,n+1}, \nabla \cdot \mathbf{v})_\Omega \\
&= (p^{m,n+1}, \mathbf{v} \cdot \mathbf{n})_{\partial\Omega} - (p^{m,n+1}, \nabla \cdot \mathbf{v})_\Omega \\
&= (g, \mathbf{v} \cdot \mathbf{n})_{\Gamma_D^f} - (p^{m,n+1}, \nabla \cdot \mathbf{v})_\Omega \qquad (5.6)
\end{aligned}
$$

where we invoke $\mathbf{v} \cdot \mathbf{n} = 0$ on Γ_N^f. In lieu of (5.5) and (5.6), we get

$$
(\boldsymbol{\kappa}^{-1}\mathbf{z}^{m,n+1}, \mathbf{v})_\Omega = -(g, \mathbf{v} \cdot \mathbf{n})_{\Gamma_D^f} + (p^{m,n+1}, \nabla \cdot \mathbf{v})_\Omega + (\rho_0 \mathbf{g}, \mathbf{v})_\Omega
$$

Replacing m by $m+1$ and subtracting the two equations, we get

$$
(\boldsymbol{\kappa}^{-1}\delta^{(m)}\mathbf{z}_h, \mathbf{v}_h)_\Omega = (\delta^{(m)}p_h, \nabla \cdot \mathbf{v}_h)_\Omega
$$

5.4.2 Discrete variational statement for the poromechanics subproblem

The weak form of the linear momentum balance (2.6) is given by

$$
(\nabla \cdot \boldsymbol{\sigma}, \mathbf{q})_\Omega + (\mathbf{f} \cdot \mathbf{q})_\Omega = 0 \qquad (\forall\ \mathbf{q} \in \mathbf{U}(\Omega)) \qquad (5.7)
$$

where $\mathbf{U}(\Omega)$ is given by

$$
\mathbf{U}(\Omega) \equiv \left\{ \mathbf{q} = (u, v, w) : u, v, w \in H^1(\Omega), \mathbf{q} = \mathbf{0} \text{ on } \Gamma_D^p \right\}
$$

We know from tensor calculus that

$$
(\nabla \cdot \boldsymbol{\sigma}, \mathbf{q})_\Omega \equiv (\nabla, \boldsymbol{\sigma}\mathbf{q})_\Omega - (\boldsymbol{\sigma} : \nabla \mathbf{q})_\Omega \qquad (5.8)
$$

Further, using the divergence theorem and the symmetry of $\boldsymbol{\sigma}$, we arrive at

$$(\nabla, \boldsymbol{\sigma}\mathbf{q})_\Omega \equiv (\mathbf{q}, \boldsymbol{\sigma}\mathbf{n})_{\partial\Omega} \tag{5.9}$$

We decompose $\nabla\mathbf{q}$ into a symmetric part $(\nabla\mathbf{q})_s \equiv \frac{1}{2}\big(\nabla\mathbf{q} + (\nabla\mathbf{q})^T\big) \equiv \boldsymbol{\epsilon}(\mathbf{q})$ and skew-symmetric part $(\nabla\mathbf{q})_{ss}$ and note that the contraction between a symmetric and skew-symmetric tensor is zero to obtain

$$\boldsymbol{\sigma} : \nabla\mathbf{q} \equiv \boldsymbol{\sigma} : (\nabla\mathbf{q})_s + \cancelto{0}{\boldsymbol{\sigma} : (\nabla\mathbf{q})_{ss}} = \boldsymbol{\sigma} : \boldsymbol{\epsilon}(\mathbf{q}) \tag{5.10}$$

From (5.7), (5.8), (5.9) and (5.10), we get

$$(\boldsymbol{\sigma}\mathbf{n}, \mathbf{q})_{\partial\Omega} - (\boldsymbol{\sigma} : \boldsymbol{\epsilon}(\mathbf{q}))_\Omega + (\mathbf{f}, \mathbf{q})_\Omega = 0$$

which, after invoking the traction boundary condition, results in the discrete weak form for the m^{th} coupling iteration as

$$(\mathbf{t}^{n+1}, \mathbf{q}_h)_{\Gamma_N^p} - (\boldsymbol{\sigma}^{m,n+1} : \boldsymbol{\epsilon}(\mathbf{q}_h))_\Omega + (\mathbf{f}^{n+1}, \mathbf{q}_h)_\Omega = 0$$

Replacing m by $m+1$ and subtracting the two equations, we get

$$(\delta^{(m)}\boldsymbol{\sigma} : \boldsymbol{\epsilon}(\mathbf{q}_h))_\Omega = 0$$

5.4.3 Summary of discrete variational statements in terms of coupling iteration differences

The discrete variational statements in terms of coupling iteration differences is : find $\delta^{(m)}p_h \in W_h$, $\delta^{(m)}\mathbf{z}_h \in \mathbf{V}_h$ and $\delta^{(m)}\mathbf{u}_h \in \mathbf{U}_h$ such that

$$C(\delta^{(m)}p_h, \theta_h)_\Omega + \Delta t(\nabla \cdot \delta^{(m)}\mathbf{z}_h, \theta_h)_\Omega = -\frac{C}{3}(\mathbf{B} : \delta^{(m-1)}\boldsymbol{\sigma}, \theta_h)_\Omega \tag{5.11}$$

$$(\boldsymbol{\kappa}^{-1}\delta^{(m)}\mathbf{z}_h, \mathbf{v}_h)_\Omega = (\delta^{(m)}p_h, \nabla\cdot\mathbf{v}_h)_\Omega \tag{5.12}$$

$$(\delta^{(m)}\boldsymbol{\sigma} : \boldsymbol{\epsilon}(\mathbf{q}_h))_\Omega = 0 \tag{5.13}$$

where the finite dimensional spaces W_h, $\mathbf{V}_h$ and $\mathbf{U}_h$ are

$$W_h = \left\{\theta_h : \theta_h| \in P_0(E)\ \forall E \in \mathscr{T}_h\right\}$$

$$\mathbf{V}_h = \left\{\mathbf{v}_h : \mathbf{v}_h|_E \leftrightarrow \hat{\mathbf{v}}|_{\hat{E}} \in \hat{\mathbf{V}}(\hat{E})\ \forall E \in \mathscr{T}_h,\ \mathbf{v}_h\cdot\mathbf{n} = 0 \text{ on } \Gamma_N^f\right\}$$

$$\mathbf{U}_h = \left\{\mathbf{q}_h = (u,v,w)|_E \in Q_1(E)\ \forall E \in \mathscr{T}_h,\ \mathbf{q}_h = \mathbf{0} \text{ on } \Gamma_D^p\right\}$$

Theorem 5.4.1. *The fixed stress split iterative coupling scheme for anisotropic poroelasticity with Biot tensor in which the flow problem is solved first by freezing all components of the stress tensor is a contraction given by*

$$\|\mathbf{B} : \delta^{(m)}\boldsymbol{\sigma}\|_\Omega^2 + \overbrace{2\|\delta^{(m)}p_h\|_\Omega^2}^{>0} + \overbrace{\frac{6\Delta t}{C}\|\boldsymbol{\kappa}^{-1/2}\delta^{(m)}\mathbf{z}_h\|_\Omega^2}^{>0} + \overbrace{\frac{6}{C}(\delta^{(m)}\boldsymbol{\sigma} : \mathbb{C}\delta^{(m)}\boldsymbol{\sigma})_\Omega}^{\geq 0}$$
$$+ \overbrace{\frac{3}{C^2}\|\delta^{(m)}\zeta\|_\Omega^2}^{>0} - \frac{6}{C^2}\|\delta^{(m)}\zeta - \delta_f^{(m)}\zeta\|_\Omega^2 \leq \|\mathbf{B} : \delta^{(m-1)}\boldsymbol{\sigma}\|_\Omega^2$$

where the quantity $\|\delta^{(m)}\zeta - \delta_f^{(m)}\zeta\|_\Omega^2$ *is driven to a small value via the convergence criterion for the algorithm.*

Proof. • **Step 1: Flow equations**

Testing (5.11) with $\theta_h \equiv \delta^{(m)}p_h$, we get

$$C\|\delta^{(m)}p_h\|_\Omega^2 + \Delta t(\nabla\cdot\delta^{(m)}\mathbf{z}_h, \delta^{(m)}p_h)_\Omega = -\frac{C}{3}(\mathbf{B} : \delta^{(m-1)}\boldsymbol{\sigma}, \delta^{(m)}p_h)_\Omega \tag{5.14}$$

Testing (5.12) with $\mathbf{v}_h \equiv \delta^{(m)}\mathbf{z}_h$, we get

$$\|\boldsymbol{\kappa}^{-1/2}\delta^{(m)}\mathbf{z}_h\|_\Omega^2 = (\delta^{(m)}p_h, \nabla\cdot\delta^{(m)}\mathbf{z}_h)_\Omega \tag{5.15}$$

From (5.14) and (5.15), we get

$$C\|\delta^{(m)}p_h\|_\Omega^2 + \Delta t\|\boldsymbol{\kappa}^{-1/2}\delta^{(m)}\mathbf{z}_h\|_\Omega^2 = -\frac{C}{3}(\mathbf{B} : \delta^{(m-1)}\boldsymbol{\sigma}, \delta^{(m)}p_h)_\Omega \tag{5.16}$$

- **Step 2: Poromechanics equations**

Testing (5.13) with $\mathbf{q}_h \equiv \delta^{(m)}\mathbf{u}_h$, we get

$$(\delta^{(m)}\boldsymbol{\sigma} : \delta^{(m)}\boldsymbol{\epsilon})_\Omega = 0 \tag{5.17}$$

We now invoke (5.2) to arrive at the expression for change in strain tensor over the $(m+1)^{th}$ coupling iteration as follows

$$\delta^{(m)}\boldsymbol{\epsilon} = \mathbb{C}\delta^{(m)}\boldsymbol{\sigma} + \frac{C}{3}\mathbf{B}\delta^{(m)}p_h \tag{5.18}$$

Substituting (5.18) in (5.17), we get

$$(\delta^{(m)}\boldsymbol{\sigma} : \mathbb{C}\delta^{(m)}\boldsymbol{\sigma})_\Omega + \frac{C}{3}(\mathbf{B} : \delta^{(m)}\boldsymbol{\sigma}, \delta^{(m)}p_h)_\Omega = 0 \tag{5.19}$$

- **Step 3: Combining flow and poromechanics equations**

Adding (5.16) and (5.19), we get

$$\begin{aligned} C\|\delta^{(m)}p_h\|_\Omega^2 + \Delta t\|\boldsymbol{\kappa}^{-1/2}\delta^{(m)}\mathbf{z}_h\|_\Omega^2 + (\delta^{(m)}\boldsymbol{\sigma} : \mathbb{C}\delta^{(m)}\boldsymbol{\sigma})_\Omega \\ + \frac{C}{3}(\mathbf{B} : \delta^{(m)}\boldsymbol{\sigma}, \delta^{(m)}p_h)_\Omega = -\frac{C}{3}(\mathbf{B} : \delta^{(m-1)}\boldsymbol{\sigma}, \delta^{(m)}p_h)_\Omega \end{aligned} \tag{5.20}$$

- **Step 4: Variation in fluid content**

In lieu of (5.3), the variation in fluid content in the $(m+1)^{th}$ coupling iteration is

$$\delta^{(m)}\zeta = C\delta^{(m)}p_h + \frac{C}{3}\mathbf{B} : \delta^{(m)}\boldsymbol{\sigma} \tag{5.21}$$

As a result, we can write

$$\frac{1}{2C}\|\delta^{(m)}\zeta\|_{\Omega}^{2} - \frac{C}{2}\|\delta^{(m)}p_h\|_{\Omega}^{2} - \frac{C}{18}\|\mathbf{B} : \delta^{(m)}\boldsymbol{\sigma}\|_{\Omega}^{2} = \frac{C}{3}(\mathbf{B} : \delta^{(m)}\boldsymbol{\sigma}, \delta^{(m)}p_h)_{\Omega} \tag{5.22}$$

From (5.20) and (5.22), we get

$$\begin{aligned} &C\|\delta^{(m)}p_h\|_{\Omega}^{2} + \Delta t\|\boldsymbol{\kappa}^{-1/2}\delta^{(m)}\mathbf{z}_h\|_{\Omega}^{2} + (\delta^{(m)}\boldsymbol{\sigma} : \mathbb{C}\delta^{(m)}\boldsymbol{\sigma})_{\Omega} + \frac{1}{2C}\|\delta^{(m)}\zeta\|_{\Omega}^{2} \\ &- \frac{C}{2}\|\delta^{(m)}p_h\|_{\Omega}^{2} - \frac{C}{18}\|\mathbf{B} : \delta^{(m)}\boldsymbol{\sigma}\|_{\Omega}^{2} = -\frac{C}{3}(\mathbf{B} : \delta^{(m-1)}\boldsymbol{\sigma}, \delta^{(m)}p_h)_{\Omega} \end{aligned} \tag{5.23}$$

Adding and subtracting $\frac{C}{6}\|\mathbf{B} : \delta^{(m)}\boldsymbol{\sigma}\|_{\Omega}^{2}$ to the LHS of (5.23) results in

$$\begin{aligned} &\frac{C}{6}\|\mathbf{B} : \delta^{(m)}\boldsymbol{\sigma}\|_{\Omega}^{2} + \frac{C}{2}\|\delta^{(m)}p_h\|_{\Omega}^{2} + \Delta t\|\boldsymbol{\kappa}^{-1/2}\delta^{(m)}\mathbf{z}_h\|_{\Omega}^{2} + (\delta^{(m)}\boldsymbol{\sigma} : \mathbb{C}\delta^{(m)}\boldsymbol{\sigma})_{\Omega} \\ &+ \frac{1}{2C}\|\delta^{(m)}\zeta\|_{\Omega}^{2} - \frac{C}{9}\|\mathbf{B} : \delta^{(m)}\boldsymbol{\sigma}\|_{\Omega}^{2} = -\frac{C}{3}(\mathbf{B} : \delta^{(m-1)}\boldsymbol{\sigma}, \delta^{(m)}p_h)_{\Omega} \end{aligned} \tag{5.24}$$

In lieu of (5.3) and the fixed stress constraint during the flow solve, the variation in fluid content during the flow solve in the $(m+1)^{th}$ coupling iteration is given by

$$\delta_f^{(m)}\zeta = C\delta_f^{(m)}p_h + \frac{C}{3}\mathbf{B} : \cancelto{\mathbf{0}}{\delta_f^{(m)}\boldsymbol{\sigma}}$$

Further, since the pore pressure is frozen during the poromechanical solve, we have $\delta_f^{(m)}p_h = \delta^{(m)}p_h$. As a result, we can write

$$\delta_f^{(m)}\zeta = C\delta^{(m)}p_h \tag{5.25}$$

Subtracting (5.25) from (5.21), we can write

$$\delta^{(m)}\zeta - \delta_f^{(m)}\zeta = \frac{C}{3}\mathbf{B} : \delta^{(m)}\boldsymbol{\sigma}$$

which implies that

$$\frac{1}{C}\|\delta^{(m)}\zeta - \delta_f^{(m)}\zeta\|_\Omega^2 = \frac{C}{9}\|\mathbf{B} : \delta^{(m)}\boldsymbol{\sigma}\|_\Omega^2 \tag{5.26}$$

In lieu of (5.26), we can write (5.24) as

$$\begin{aligned}&\frac{C}{6}\|\mathbf{B} : \delta^{(m)}\boldsymbol{\sigma}\|_\Omega^2 + \frac{C}{2}\|\delta^{(m)}p_h\|_\Omega^2 + \Delta t\|\boldsymbol{\kappa}^{-1/2}\delta^{(m)}\mathbf{z}_h\|_\Omega^2 + (\delta^{(m)}\boldsymbol{\sigma} : \mathbb{C}\delta^{(m)}\boldsymbol{\sigma})_\Omega \\ &+ \frac{1}{2C}\|\delta^{(m)}\zeta\|_\Omega^2 - \frac{1}{C}\|\delta^{(m)}\zeta - \delta_f^{(m)}\zeta\|_\Omega^2 = -\frac{C}{3}(\mathbf{B} : \delta^{(m-1)}\boldsymbol{\sigma}, \delta^{(m)}p_h)_\Omega\end{aligned} \tag{5.27}$$

• Step 5: Invoking positive-semidefiniteness of the compliance tensor

The fourth order anisotropic compliance tensor $\mathbb{C}$ is positive-semidefinite (see Gurtin et al. [41]) in the sense that it obeys $\mathbf{S} : \mathbb{C}\mathbf{S} \geq 0$ for all symmetric tensors $\mathbf{S}$. In lieu of the above, since the Cauchy stress tensor is symmetric, the following holds true

$$\delta^{(m)}\boldsymbol{\sigma} : \mathbb{C}\delta^{(m)}\boldsymbol{\sigma} \geq 0 \tag{5.28}$$

In lieu of (5.28), we can write (5.27) as

$$\begin{aligned}&\frac{C}{6}\|\mathbf{B} : \delta^{(m)}\boldsymbol{\sigma}\|_\Omega^2 + \overbrace{\frac{C}{2}\|\delta^{(m)}p_h\|_\Omega^2}^{>0} + \overbrace{\Delta t\|\boldsymbol{\kappa}^{-1/2}\delta^{(m)}\mathbf{z}_h\|_\Omega^2}^{>0} + \overbrace{(\delta^{(m)}\boldsymbol{\sigma} : \mathbb{C}\delta^{(m)}\boldsymbol{\sigma})_\Omega}^{\geq 0} \\ &+ \overbrace{\frac{1}{2C}\|\delta^{(m)}\zeta\|_\Omega^2}^{>0} - \frac{1}{C}\|\delta^{(m)}\zeta - \delta_f^{(m)}\zeta\|_\Omega^2 = -\frac{C}{3}(\mathbf{B} : \delta^{(m-1)}\boldsymbol{\sigma}, \delta^{(m)}p_h)_\Omega\end{aligned} \tag{5.29}$$

The quantity $\|\delta^{(m)}\zeta - \delta_f^{(m)}\zeta\|_\Omega^2$ on the LHS of (5.29) is driven to a small value via the convergence criterion for the algorithm as explained in Section 5.5.

• Step 6: Invoking the Young's inequality

Since the sum of the terms on the LHS of (5.29) is nonnegative, the RHS is also nonnegative. We invoke the Young's inequality (see Steele [82])

$$|ab| \leq \frac{a^2}{2} + \frac{b^2}{2} \qquad (\forall\ a, b \in \mathbb{R})$$

for the RHS of (5.29) as follows

$$-\frac{C}{3}(\mathbf{B} : \delta^{(m-1)}\boldsymbol{\sigma}, \delta^{(m)}p_h)_\Omega \leq \frac{C}{3}\left(\frac{1}{2}\|\mathbf{B} : \delta^{(m-1)}\boldsymbol{\sigma}\|^2_\Omega + \frac{1}{2}\|\delta^{(m)}p_h\|^2_\Omega\right) \tag{5.30}$$

In lieu of (5.30), we write (5.29) as

$$\frac{C}{6}\|\mathbf{B} : \delta^{(m)}\boldsymbol{\sigma}\|^2_\Omega + \overbrace{\frac{C}{2}\|\delta^{(m)}p_h\|^2_\Omega}^{>0} + \overbrace{\Delta t\|\boldsymbol{\kappa}^{-1/2}\delta^{(m)}\mathbf{z}_h\|^2_\Omega}^{>0} + \overbrace{(\delta^{(m)}\boldsymbol{\sigma} : \mathbb{C}\delta^{(m)}\boldsymbol{\sigma})_\Omega}^{\geq 0}$$

$$+ \overbrace{\frac{1}{2C}\|\delta^{(m)}\zeta\|^2_\Omega}^{>0} - \frac{1}{C}\|\delta^{(m)}\zeta - \delta^{(m)}_f\zeta\|^2_\Omega \leq \frac{C}{6}\|\mathbf{B} : \delta^{(m-1)}\boldsymbol{\sigma}\|^2_\Omega + \frac{C}{6}\|\delta^{(m)}p_h\|^2_\Omega$$

which can also be written as

$$\|\mathbf{B} : \delta^{(m)}\boldsymbol{\sigma}\|^2_\Omega + \overbrace{2\|\delta^{(m)}p_h\|^2_\Omega}^{>0} + \overbrace{\frac{6\Delta t}{C}\|\boldsymbol{\kappa}^{-1/2}\delta^{(m)}\mathbf{z}_h\|^2_\Omega}^{>0} + \overbrace{\frac{6}{C}(\delta^{(m)}\boldsymbol{\sigma} : \mathbb{C}\delta^{(m)}\boldsymbol{\sigma})_\Omega}^{\geq 0}$$

$$+ \overbrace{\frac{3}{C^2}\|\delta^{(m)}\zeta\|^2_\Omega}^{>0} - \frac{6}{C^2}\|\delta^{(m)}\zeta - \delta^{(m)}_f\zeta\|^2_\Omega \leq \|\mathbf{B} : \delta^{(m-1)}\boldsymbol{\sigma}\|^2_\Omega \tag{5.31}$$

□

5.5 Convergence criterion

In lieu of (5.3), the variation in fluid content $\delta^{(m)}_f\zeta$ measured during the flow solve in the $(m+1)^{th}$ coupling iteration at any time step is

$$\delta^{(m)}_f\zeta = \frac{1}{M}\delta^{(m)}_f p_h + \boldsymbol{\alpha} : \delta^{(m)}_f\boldsymbol{\epsilon} = \zeta^{m+1}_p - \zeta^m \tag{5.32}$$

where ζ^m is the fluid content at the end of the previous or m^{th} coupling iteration and ζ_p^{m+1} serves as the predictor to the fluid content at the end of the current or $(m+1)^{th}$ coupling iteration. Similarly, the variation in fluid content $\delta^{(m)}\zeta$ over the $(m+1)^{th}$ coupling iteration (including the flow solve and poromechanics solve) at any time step is

$$\delta^{(m)}\zeta = \frac{1}{M}\delta^{(m)}p_h + \boldsymbol{\alpha} : \delta^{(m)}\boldsymbol{\epsilon} = \zeta_c^{m+1} - \zeta^m \tag{5.33}$$

where ζ_c^{m+1} serves as the corrector to ζ_p^{m+1}. Subtracting (5.32) from (5.33), we get

$$\delta^{(m)}\zeta - \delta_f^{(m)}\zeta \equiv \zeta_c^{m+1} - \zeta_p^{m+1} = \boldsymbol{\alpha} : (\delta^{(m)}\boldsymbol{\epsilon} - \delta_f^{(m)}\boldsymbol{\epsilon}) = \boldsymbol{\alpha} : (\boldsymbol{\epsilon}_c^{m+1} - \boldsymbol{\epsilon}_p^{m+1})$$

which implies that the difference between the predicted value and the corrected value of the fluid content at the end of the $(m+1)^{th}$ coupling iteration is equal to the difference between the predicted value ($\boldsymbol{\epsilon}_p^{m+1}$) and the corrected value ($\boldsymbol{\epsilon}_c^{m+1}$) of the strain tensor at the end of the $(m+1)^{th}$ coupling iteration scaled by the Biot tensor $\boldsymbol{\alpha}$.

In lieu of the above, the stopping criterion for coupling iterations at any time step is

$$\left\| \frac{\delta^{(m)}\zeta - \delta_f^{(m)}\zeta}{\zeta_c^{m+1}} \right\|_{L^\infty} \equiv \left\| \frac{\zeta_c^{m+1} - \zeta_p^{m+1}}{\zeta_c^{m+1}} \right\|_{L^\infty} \leq TOL$$

where $TOL > 0$ is a pre-specified tolerance. We set $TOL = 1 \times 10^{-8}$ for our simulations.

5.6 Contracted notation, material symmetry and transverse isotropy

The generalized Hooke's law (5.1) written in indicial notation

$$\sigma_{ij} = \mathbb{M}_{ijkl}\epsilon_{kl} - \alpha_{ij}p$$

is rewritten in contracted notation as

$$\sigma_\beta = \mathbb{M}_{\beta\gamma}\epsilon_\gamma - \alpha_\beta p$$

where the transformation is accomplished by replacing the subscripts ij (or kl) by β (or γ) using the following rules

ij (or kl)		β (or γ)
11	$\longleftrightarrow$	1
22	$\longleftrightarrow$	2
33	$\longleftrightarrow$	3
23 (or 32)	$\longleftrightarrow$	4
31 (or 13)	$\longleftrightarrow$	5
12 (or 21)	$\longleftrightarrow$	6

In other words, the stress and strain tensors are represented as

$$\boldsymbol{\sigma} = \left\{ \begin{array}{cccccc} \sigma_{xx} & \sigma_{yy} & \sigma_{zz} & \sigma_{yz} & \sigma_{xz} & \sigma_{xy} \end{array} \right\}^T,$$
$$\boldsymbol{\epsilon} = \left\{ \begin{array}{cccccc} \epsilon_{xx} & \epsilon_{yy} & \epsilon_{zz} & \epsilon_{yz} & \epsilon_{xz} & \epsilon_{xy} \end{array} \right\}^T$$

A material is said to possess a symmetry with respect to an orthogonal transformation $\boldsymbol{\chi}$ if the elasticity tensor $\mathbb{M}$ is invariant under the orthogonal transformation $\boldsymbol{\chi}$ as follows (see Ting [87])

$$\mathbb{M}' = \mathbf{K}\mathbb{M}\mathbf{K}^T \equiv \mathbb{M}$$

where $\mathbb{M}'$ is the transformed elasticity tensor and $\mathbf{K}$ is given by

$$\mathbf{K} = \begin{bmatrix} \chi_{11}^2 & \chi_{12}^2 & \chi_{13}^2 & 2\chi_{12}\chi_{13} & 2\chi_{13}\chi_{11} & 2\chi_{11}\chi_{12} \\ \chi_{21}^2 & \chi_{22}^2 & \chi_{23}^2 & 2\chi_{22}\chi_{23} & 2\chi_{23}\chi_{21} & 2\chi_{21}\chi_{22} \\ \chi_{31}^2 & \chi_{32}^2 & \chi_{33}^2 & 2\chi_{32}\chi_{33} & 2\chi_{33}\chi_{31} & 2\chi_{31}\chi_{32} \\ \chi_{21}\chi_{31} & \chi_{22}\chi_{32} & \chi_{23}\chi_{33} & & & \\ \chi_{31}\chi_{11} & \chi_{32}\chi_{12} & \chi_{33}\chi_{13} & & \mathbf{K}' & \\ \chi_{11}\chi_{21} & \chi_{12}\chi_{22} & \chi_{13}\chi_{23} & & & \end{bmatrix}$$

$$\mathbf{K}' = \begin{bmatrix} \chi_{22}\chi_{33} + \chi_{23}\chi_{32} & \chi_{23}\chi_{31} + \chi_{21}\chi_{33} & \chi_{21}\chi_{32} + \chi_{22}\chi_{31} \\ \chi_{32}\chi_{13} + \chi_{33}\chi_{12} & \chi_{33}\chi_{11} + \chi_{31}\chi_{13} & \chi_{31}\chi_{12} + \chi_{32}\chi_{11} \\ \chi_{12}\chi_{23} + \chi_{13}\chi_{22} & \chi_{13}\chi_{21} + \chi_{11}\chi_{23} & \chi_{11}\chi_{22} + \chi_{12}\chi_{21} \end{bmatrix}$$

where $\boldsymbol{\chi}$ given by

$$\boldsymbol{\chi} = \mathbf{I} - 2\mathbf{n}\mathbf{n}^T$$

is a reflection with respect to a plane whose normal is $\mathbf{n}$. The plane is also refered to as the plane of material symmetry. The symmetry planes can be one of the following three possibilities

1. $x = 0$ plane and any plane that contains the x-axis
2. $y = 0$ plane and any plane that contains the y-axis
3. $z = 0$ plane and any plane that contains the z-axis

For the Mandel's problem (see Figure 5.2), the material symmetry planes are the second possibility ($y = 0$ plane and any plane that contains the y-axis). In lieu of that, the elasticity tensor $\mathbb{M}$ reduces to

$$\begin{bmatrix} \mathbb{M}_{11} & \mathbb{M}_{12} & \mathbb{M}_{13} & 0 & 0 & 0 \\ \mathbb{M}_{12} & \mathbb{M}_{22} & \mathbb{M}_{12} & 0 & 0 & 0 \\ \mathbb{M}_{13} & \mathbb{M}_{12} & \mathbb{M}_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & 2\mathbb{M}_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & (\mathbb{M}_{11} - \mathbb{M}_{13}) & 0 \\ 0 & 0 & 0 & 0 & 0 & 2\mathbb{M}_{44} \end{bmatrix} \tag{5.34}$$

with five independent componenents $\mathbb{M}_{11}$, $\mathbb{M}_{12}$, $\mathbb{M}_{13}$, $\mathbb{M}_{22}$ and $\mathbb{M}_{44}$.

5.7 Mandel's problem

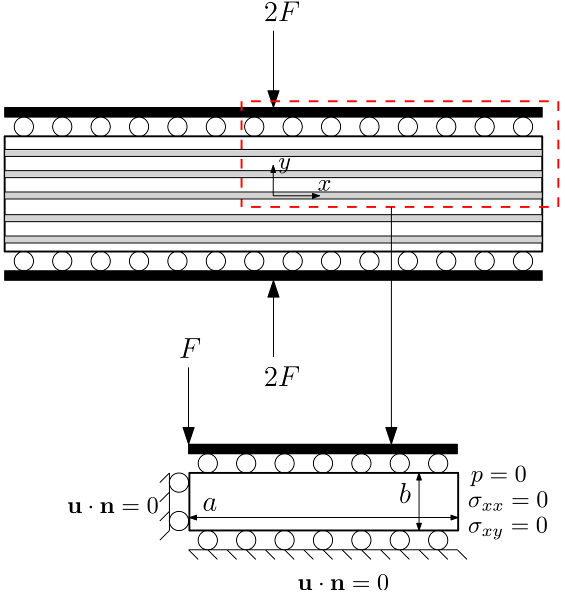

Figure 5.2: Circles indicate rollers and solid black boxes indicate rigid frictionless plates. The material symmetry planes are the $y = 0$ plane and any plane that contains the y-axis.

The algorithm is implemented in the in-house parallel reservoir simulator at the Center for Subsurface Modeling. The bevo3 supercomputer at the Institute for Computational Engineering and Sciences is employed for all the simulations. The analytical solution provided by Abousleiman et al. [1] to the Mandel's problem for transversely isotropic poroelasticity with compressible fluid and solid components serves as a benchmark for validation of coupled flow and poroelasticity codes.

5.7.1 Geometry

As shown in Figure 5.2, an infinitely long rectangular specimen is sandwiched between rigid, frictionless plates. For the transverse isotropy, the y-axis is assigned as the axis of material symmetry, which implies that at every point of the specimen along the xz-plane, the mechanical properties are the same in all directions. The lateral sides are free from normal and shear stress and pore pressure. At $t = 0^+$, a force intensity of $2F\ N/m$ is applied to the rigid plates.

5.7.2 Initial and boundary conditions

The initial and boundary conditions are

$$\sigma_{xx}|_{t=0} = \sigma_{xy}|_{t=0} = \sigma_{yy}|_{t=0} = 0, p|_{t=0} = 0 \qquad (\forall x, y)$$

$$\sigma_{xx}|_{x=\pm a} = \sigma_{xy}|_{x=\pm a} = \sigma_{yx}|_{y=\pm b} = p|_{x=\pm a} = 0 \qquad (\forall t)$$

$$\int\limits_{-a}^{a} \sigma_{yy} dx = -2F,\ \mathbf{u}\cdot\mathbf{n} = 0 \qquad (\forall t, y = \pm b)$$

where $\mathbf{n}$ is unit outward normal to the boundary. Plane strain condition is applicable i.e. $\epsilon_{zz} = 0$. Given the biaxial symmetry of the problem, only a quarter of the domain needs to be modeled as shown in Figure 5.2. Following the approach of Mikelić et al. [60], the boundary conditions are recast as

$$\sigma_{xx}|_{x=a} = \sigma_{xy}|_{x=a} = \sigma_{yx}|_{y=b} = 0, \mathbf{u}\cdot\mathbf{n}|_{y=b} = U_y \qquad (\forall t)$$

$$\mathbf{u}\cdot\mathbf{n}|_{x=0} = \mathbf{u}\cdot\mathbf{n}|_{y=0} = 0,\ p|_{x=a} = 0 \qquad (\forall t)$$

$$\mathbf{u}\cdot\mathbf{n}|_{x=0} = \mathbf{u}\cdot\mathbf{n}|_{y=0} = \mathbf{u}\cdot\mathbf{n}|_{y=b} = 0 \qquad (\forall t)$$

where U_y represents the analytical solution for u_y at $y = b$.

5.7.3 Stress-strain relations

In lieu of (5.34), the stress-strain relations (with the plane strain assumption $\epsilon_{zz} = 0$ and $\epsilon_{yz} = \epsilon_{xz} = 0$) are

$$\begin{Bmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \sigma_{zz} \\ \sigma_{yz} \\ \sigma_{xz} \\ \sigma_{xy} \end{Bmatrix} = \begin{bmatrix} \mathbb{M}_{11} & \mathbb{M}_{12} & \mathbb{M}_{13} & 0 & 0 & 0 \\ \mathbb{M}_{12} & \mathbb{M}_{22} & \mathbb{M}_{12} & 0 & 0 & 0 \\ \mathbb{M}_{13} & \mathbb{M}_{12} & \mathbb{M}_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & 2\mathbb{M}_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & 2\mathbb{M}_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & 2\mathbb{M}_{44} \end{bmatrix} \begin{Bmatrix} \epsilon_{xx} \\ \epsilon_{yy} \\ \cancelto{0}{\epsilon_{zz}} \\ \cancelto{0}{\epsilon_{yz}} \\ \cancelto{0}{\epsilon_{xz}} \\ \epsilon_{xy} \end{Bmatrix}$$
$$- \begin{bmatrix} \alpha_{xx} & 0 & 0 & 0 & 0 & 0 \\ 0 & \alpha_{yy} & 0 & 0 & 0 & 0 \\ 0 & 0 & \alpha_{xx} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} p$$

where $\mathbb{M}_{11}$, $\mathbb{M}_{12}$, $\mathbb{M}_{13}$, $\mathbb{M}_{22}$, $\mathbb{M}_{44}$ and $\mathbb{M}_{55}$ are given by (see Boresi et al. [13])

$$\mathbb{M}_{11} = \frac{E_x(E_y - E_x\nu_{yx}^2)}{(1+\nu_{zx})(E_y - E_y\nu_{zx} - 2E_x\nu_{yx}^2)} \tag{5.35a}$$

$$\mathbb{M}_{12} = \frac{E_xE_y\nu_{yx}}{E_y - E_y\nu_{zx} - 2E_x\nu_{yx}^2} \tag{5.35b}$$

$$\mathbb{M}_{13} = \frac{E_x(\nu_{zx}E_y + E_x\nu_{yx}^2)}{(1+\nu_{zx})(E_y - E_y\nu_{zx} - 2E_x\nu_{yx}^2)} \tag{5.35c}$$

$$\mathbb{M}_{22} = \frac{E_y^2(1-\nu_{zx})}{E_y - E_y\nu_{zx} - 2E_x\nu_{yx}^2} \tag{5.35d}$$

$$\mathbb{M}_{44} = G_{yz} = \frac{E_y}{2(1+\nu_{yx})} \equiv \frac{E_x}{2(1+\nu_{xy})} \tag{5.35e}$$

$$\mathbb{M}_{55} = (\mathbb{M}_{11} - \mathbb{M}_{13})/2 \tag{5.35f}$$

The set of data reported in Aoki et al. [4] for Trafalgar shale are employed here

$$E_x = 20.6\, GPa;\ E_y = 17.3\, GPa;\ \nu_{zx} = 0.189;\ \nu_{yx} = 0.246$$

As a result, the five independent material parameters ((5.35a)-(5.35f)) are obtained as

$$\mathbb{M}_{11} = 24.1\, GPa;\ \mathbb{M}_{12} = 7.62\, GPa;\ \mathbb{M}_{13} = 6.8\, GPa;$$

$$\mathbb{M}_{22} = 21.0\, GPa;\ \mathbb{M}_{44} = 6.94\, GPa;\ \mathbb{M}_{55} = 8.66\, GPa$$

As a result, the elasticity matrix (see (5.34)) can be written as (all components in GPa)

$$\mathbb{M} = \begin{bmatrix} 24.1 & 7.62 & 6.8 & 0 & 0 & 0 \\ 7.62 & 21.0 & 7.62 & 0 & 0 & 0 \\ 6.8 & 7.62 & 24.1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 13.88 & 0 & 0 \\ 0 & 0 & 0 & 0 & 17.33 & 0 \\ 0 & 0 & 0 & 0 & 0 & 13.88 \end{bmatrix}$$

The compliance tensor $\mathbb{C}$ is obtained as the inverse of the elasticity tensor $\mathbb{M}$ and is given by (all components in $(GPa)^{-1}$)

$$\mathbb{C} = \begin{bmatrix} 0.0485 & -0.0142 & -0.0092 & 0 & 0 & 0 \\ -0.0142 & 0.05780 & -0.0142 & 0 & 0 & 0 \\ -0.0092 & -0.0142 & 0.0485 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.07202 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.05772 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.07202 \end{bmatrix}$$

The specimen dimensions are taken to be $a \times b \equiv 100\, m \times 10\, m$ and the point load intensity is taken to be $10^8\, N/m$. The initial porosity is taken to be 0.2, the fluid compressibility is $3.03 \times 10^{-10}\, Pa^{-1}$, the x-permeability is $K_{xx} = 100\, mD$ and fluid viscosity is $\mu = 0.001\, Pa - s$. Gravity is not active.

5.7.4 Analytical solution

The analytical solution u_x, u_y and p for the x-displacement, y-displacement and pressure respectively, are given by

$$
\begin{aligned}
u_x = & \left[\frac{F}{a} \frac{\mathbb{M}_{12}}{(\mathbb{M}_{11}\mathbb{M}_{22} - \mathbb{M}_{12}^2)} - \frac{2F}{a} \frac{\alpha_{xx}\alpha_{yy}M + \mathbb{M}_{12}}{A_1 M(\alpha_{yy}\mathbb{M}_{11} - \alpha_{xx}\mathbb{M}_{12})} \right. \\
& \left. \times \sum_{i=1}^{\infty} \frac{sin\beta_i\, cos\beta_i}{\beta_i - sin\beta_i\, cos\beta_i} e^{-\beta_i^2 c_1 t/a^2} \right] x \\
& + \frac{2F\alpha_{xx}}{A_2\mathbb{M}_{11}} \sum_{i=1}^{\infty} \frac{cos\beta_i\, sin(\beta_i x/a)}{\beta_i - sin\beta_i\, cos\beta_i} e^{-\beta_i^2 c_1 t/a^2} \\
u_y = & -\frac{F}{a} \frac{\mathbb{M}_{11}}{\mathbb{M}_{11}\mathbb{M}_{22} - \mathbb{M}_{12}^2} \\
& \times \left[1 + 2(A_2/A_1 - 1) \sum_{i=1}^{\infty} \frac{sin\beta_i\, cos\beta_i}{\beta_i - sin\beta_i\, cos\beta_i} e^{-\beta_i^2 c_1 t/a^2} \right] y \\
p = & \frac{2F}{aA_1} \sum_{i=1}^{\infty} \frac{sin\beta_i}{\beta_i - sin\beta_i\, cos\beta_i} (cos(\beta_i x/a) - cos\beta_i) e^{-\beta_i^2 c_1 t/a^2}
\end{aligned}
$$

5.7.5 Micro-homogeneity and micro-isotropy

The expressions (5.36) for α_{xx} and α_{yy} are obtained under the assumptions of micro-homogeneity and micro-isotropy. The micro-homogeneity assumption states that the skeleton of the porous material is homogeneous at the pore scale but can be heterogeneous at the macroscopic scale by having different micro-homogeneous material distributed in space (Cheng [23]). The micro-isotropy assumption states that the material is isotropic at the pore level such that it is characterized by the bulk modulus K_s (Cui et al. [26]) and the macroscopic anisotropy is the manifestation of the directionality of the skeleton or the pore structure, and not the solid constituent itself.

$$\alpha_{xx} = 1 - \frac{\mathbb{M}_{11} + \mathbb{M}_{12} + \mathbb{M}_{13}}{3K_s};\ \alpha_{yy} = 1 - \frac{2\mathbb{M}_{12} + \mathbb{M}_{22}}{3K_s} \tag{5.36}$$

Furthermore, the Biot modulus M is given by

$$M = \frac{K_s^2}{K_s(1 + \phi_0(cK_s - 1)) - \frac{2\mathbb{M}_{11}+\mathbb{M}_{22}+2\mathbb{M}_{13}+4\mathbb{M}_{12}}{9}}$$

while β_i and c_1 are given as

$$\begin{aligned} A_1 &= \frac{M(\alpha_{xx}^2\mathbb{M}_{22} - 2\alpha_{xx}\alpha_{yy}\mathbb{M}_{12} + \alpha_{yy}^2\mathbb{M}_{11}) + \mathbb{M}_{11}\mathbb{M}_{22} - \mathbb{M}_{12}^2}{M(\alpha_{yy}\mathbb{M}_{11} - \alpha_{xx}\mathbb{M}_{12})} \\ A_2 &= \frac{\alpha_{yy}\mathbb{M}_{11} - \alpha_{xx}\mathbb{M}_{12}}{\mathbb{M}_{11}};\ \frac{tan(\beta_i)}{\beta_i} = \frac{A_1}{A_2} \\ c_1 &= \frac{\kappa_{11}M\mathbb{M}_{11}}{\mathbb{M}_{11} + \alpha_{xx}^2 M} \end{aligned}$$

where $K_s = 48.2\,GPa$ is the value as employed by Abousleiman et al. [1]. As a result, we have

$$\begin{aligned} &\alpha_{xx} = 0.733;\ \alpha_{yy} = 0.749;\ M = 13.93\,GPa;\ A_1 = 3.9050; \\ &A_2 = 0.5181;\ \frac{tan(\beta_i)}{\beta_i} = 7.5372;\ c_1 = 1.0486\,m^2/s \end{aligned}$$

Furthermore, the Hooke's law constant C in (5.2) and (5.3) is given as

$$C = 2\mathbb{C}_{11} + \mathbb{C}_{22} + 2\mathbb{C}_{13} + 4\mathbb{C}_{12} - \frac{1}{K_s}(1 - \phi_0(cK_s - 1)) = 0.1154 \times 10^{-9}Pa^{-1}$$

5.7.6 Fluid content updates

The fixed stress constraint $\delta_f^{(m)}\boldsymbol{\sigma} = \mathbf{0}$ makes it convenient to measure $\delta_f^{(m)}\zeta$ using the definition of ζ in (5.3) which involves the pressure and stress as follows

$$\delta_f^{(m)}\zeta \equiv \zeta_p^{m+1} - \zeta^m = C\delta_f^{(m)} p_h + \frac{C}{3}\mathbf{B} : \cancelto{\mathbf{0}}{\delta_f^{(m)}\boldsymbol{\sigma}}$$

Since $\delta_f^{(m)} p_h$ is what we solve for in the flow system, $\delta_f^{(m)}\zeta$ can be easily measured using the above equation rather than (5.32). Since $\delta^{(m)}\boldsymbol{\epsilon}$ is obtained by post-processing the poromechanical solution, $\delta^{(m)}\zeta$ can be obtained using (5.33) as follows

$$\delta^{(m)}\zeta \equiv \zeta_c^{m+1} - \zeta^m = \frac{1}{M}\delta^{(m)} p_h + \boldsymbol{\alpha} : \delta^{(m)}\boldsymbol{\epsilon} = \frac{1}{M}\delta^{(m)} p_h + \alpha_{xx}\delta^{(m)}\epsilon_{xx} + \alpha_{yy}\delta^{(m)}\epsilon_{yy}$$

5.7.7 Optimal norm

Phillips and Wheeler [68] have shown that when mixed finite element spaces are employed for fluxes and pressures and a continuous Galerkin space is employed for displacements, optimality should be achieved when the pressure solution error is measured in the $L^\infty(L^2)$ norm. Although the analysis of Phillips and Wheeler [68] stands for the case of isotropic poroelasticity with scalar Biot parameter, intuitively, we expect a similar result for the case of anisotropic poroelasticity with tensor Biot parameter. The error norm is computed using the midpoint quadrature rule:

$$\|\frac{1}{M}(p - p_h)\|_{L^\infty(L^2)} \equiv \max_{0<\tau\leq T}\left(\sum_{E\in\mathcal{T}_h} |E|\Big(\frac{p(\tau, m_e) - p_h(m_e)}{M}\Big)^2\right)^{\frac{1}{2}}$$

where m_e is the center of mass of element E and T is the total time. To minimize the effects of the error produced by time discretization, a small time step of $1 \times 10^{-2}\,s$ is chosen. As shown in Table 5.1, we observe first order convergence for the pore pressure solution.

$\mathcal{T}_h$	$\lVert \frac{1}{M}(p - p_h) \rVert_{L^\infty(L^2)}$	Rate
10×10	0.2544×10^{-1}	
20×20	0.1355×10^{-1}	0.9088
30×30	0.8466×10^{-2}	1.1600
40×40	0.6311×10^{-2}	1.0211
50×50	0.5067×10^{-2}	0.9839

Table 5.1: Order of convergence of pore pressure solution using the fixed stress split scheme for the Mandel's problem with transverse isotropy

5.8 Summary

The work in previous chapters assumes isotropy of the porous solid. In this chapter, we again invoke the concept of contraction mapping to establish the theoretical convergence of the staggered solution algorithm for the case of linearized single phase flow coupled with linear poromechanics with flow and poromechanics being resolved on the same grid and the anisotropy of the porous rock factored in. We again establish the numerical convergence of this solution algorithm for the benchmark problem.

Chapter 6

Augmented Lagrangian for treatment of hanging nodes in the mechanics hexahedral mesh

The surge of activity in the resolution of fine scale features in the field of earth sciences over the past decade necessitates the development of robust yet simple algorithms that can tackle the various drawbacks of in silico models developed hitherto. One such drawback is that of the restrictive computational cost of finite element method in rendering resolutions to the fine scale features, while at the same time keeping the domain being modeled sufficiently large. We propose the use of the augmented Lagrangian method commonly used in the treatment of hanging nodes in contact mechanics in tackling the drawback. An interface is introduced in a typical finite element mesh across which an aggressive coarsening of the finite elements is possible. The method is based upon minimizing an augmented potential energy which factors in the constraint that exists at the hanging nodes on that interface. This allows for a significant reduction in the number of finite elements comprising the mesh with concomitant reduction in the computational expense.

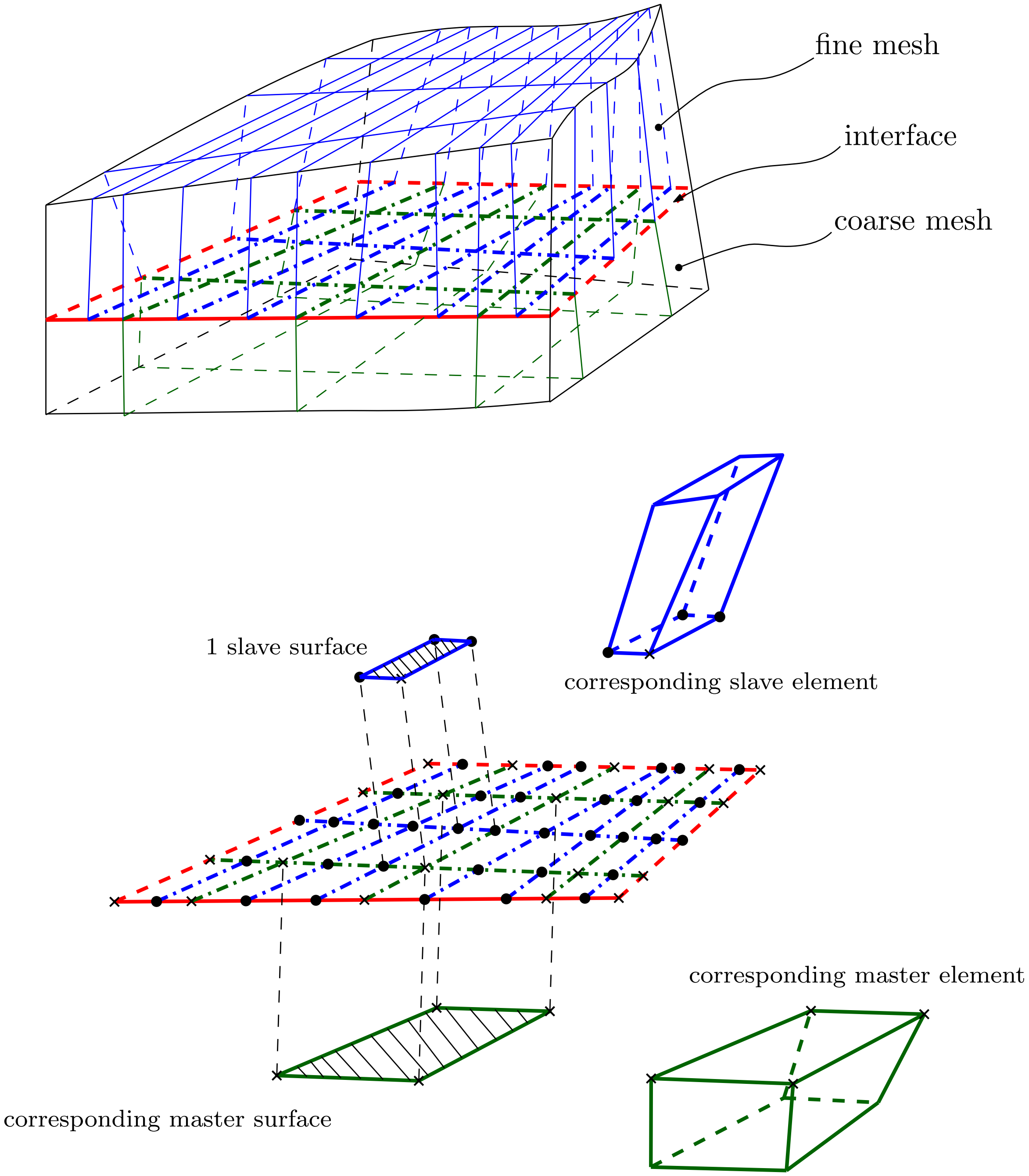

Figure 6.1: The hanging nodes are represented by dots. The boundary of the interface is represented by a red dashed line. Crosses represent the nodes that are common to both the fine and coarse mesh

The quantum of work devoted to modeling of fine scale features in the subsurface in the recent decade has spawned a need for simple yet powerful algorithms to simulate the same in silico with low computational cost. The main barrier to these simulations lies in the restrictively fine mesh that needs to be invoked to resolve the finer features of the corresponding physics, while at the same time keeping the domain under consideration sufficiently large. The most logical approach to this problem is to allow for a fine mesh to exist in the regions which need a fine mesh, and a coarse mesh to exist in regions which do not need a fine mesh. The authors have in the past developed a method to simulate subsurface flow on a fine mesh and subsurface mechanics on a coarse mesh while allowing for the coupling between the physics of flow and mechanics via a staggered solution algorithm (see Chapter 3). The aforementioned work, though, is restrictive in the sense that the mesh for the mechanics domain needs to be uniformly coarser than the mesh for the flow domain. This makes the algorithm infeasible for problems involving fine scale features for the mechanics. With that in mind, in this work, we propose an addendum to the aforementioned algorithm. We invoke the concept of hanging nodes in finite elements. It essentially means that there is an interface in the mechanics mesh across which an aggressive refinement is possible, thus allowing for fine elements on one side of the interface and coarser elements on the other side of the interface. An example is given in Figure 6.1.

6.1 Summary of the various formulations for treatment of hanging nodes developed hitherto

The concept of hanging nodes itself is not new, and has been given its due diligence as far back as in the 1980s, 1990s and early 2000s (see the works of Felippa [30], Powell [70], Hallquist et al. [42], Simo et al. [78], Wriggers and Simo [96], Parisch [66], Papadopoulos and Taylor [64], Papadopoulos and Taylor [65], McDevitt and Laursen [57], El-Abbasi and Bathe [29], Becker et al. [6], Puso and Laursen [71], Puso and Laursen [72], Wriggers [95] and Wriggers and Zavarise [97]). The problem is looked upon as optimization of a functional with a constraint which dictates the geometry of the interface of the hanging nodes. Representing $\mathscr{C}(\mathbf{u})$ as the potential energy functional of a system with $\mathbf{u}$ respresenting the displacement field, the optimization problem statement is simply put forth as: Minimize $\mathscr{C}(\mathbf{u})$ subject to a constraint $\mathbf{g}(\mathbf{u}) = \mathbf{0}$.

6.1.1 Penalty formulation

The penalty formulation penalizes the non-satisfaction of the constraint by augmenting the energy functional to be minimized as follows

$$\tilde{\mathscr{C}}(\mathbf{u}) = \mathscr{C}(\mathbf{u}) + \frac{\epsilon}{2}\mathbf{g}(\mathbf{u}) \cdot \mathbf{g}(\mathbf{u})$$

where ϵ is a large penalty parameter. A large enough penalty parameter closes the gap between the solution obtained through the penalty formulation and the original minimization problem solution. On the other hand, a large penalty parameter leads to highly ill-conditioned stiffness matrix in the eventual system of equations obtained at the discrete level. As a result, the choice of penalty parameter is a compromise between solution accuracy and solution stability.

6.1.2 Lagrangian formulation

The Lagrangian formulation, on the other hand, enforces the constraint by introducing a lagrange multiplier term to the energy functional to be minimized as follows

$$\tilde{\mathscr{C}}(\mathbf{u}, \boldsymbol{\lambda}) = \mathscr{C}(\mathbf{u}) + \boldsymbol{\lambda} \cdot \mathbf{g}(\mathbf{u})$$

where $\boldsymbol{\lambda}$ is the force conjugate to the constraint $\mathbf{g}(\mathbf{u}) = 0$, and is refered to as the lagrange multiplier. Although this method allows for the exact satisfaction of the constraint, the increase in number of degrees of freedom of the original system by the number of lagrange multipliers makes the augmentation computationally expensive.

6.1.3 Perturbed Lagrangian formulation

The perturbed Lagrangian formulation circumvents this problem by introducing the following functional to be minimized

$$\tilde{\mathscr{C}}(\mathbf{u}, \boldsymbol{\lambda}) \equiv \mathscr{C}(\mathbf{u}) + \boldsymbol{\lambda} \cdot \mathbf{g}(\mathbf{u}) - \frac{1}{2\epsilon} \boldsymbol{\lambda} \cdot \boldsymbol{\lambda}$$

with the constraint

$$\mathbf{g}(\mathbf{u}) - \frac{\boldsymbol{\lambda}}{\epsilon} = \mathbf{0}$$

where ϵ is the penalty parameter. This allows for the lagrange multiplier to be posed in terms of the constraint thus negating the need to solve for the multiplier as an additional degree of freedom. This method, though, suffers from the same problem that the original penalty method suffers from, i.e. a careful compromise between accuracy and stability must be made in the choice of the penalty parameter.

6.1.4 Augmented Lagrangian formulation

The augmented Lagrangian formulation circumvents this issue by introducing the following functional to be minimized

$$\tilde{\mathscr{C}}(\mathbf{u}, \boldsymbol{\lambda}^k) \equiv \mathscr{C}(\mathbf{u}) + \boldsymbol{\lambda}^k \cdot \mathbf{g}(\mathbf{u}) + \frac{\epsilon}{2}\mathbf{g}(\mathbf{u}) \cdot \mathbf{g}(\mathbf{u})$$

with the constraint

$$\boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^k = \epsilon\mathbf{g}(\mathbf{u})$$

where $\boldsymbol{\lambda}^k$ is the lagrange multiplier evaluated at the k^{th} iteration. As is evident from the formulation, the lagrange multiplier is evaluated iteratively till it reaches an asymptotic value. The lagrange multiplier, is not an additional degree of freedom, and hence the system size does not increase as compared to the original minimization problem. The biggest advantage of this method is that the solution stability is not a function of the penalty parameter, and furthermore the lagrange multiplier iterative process reaches the true asymptotic value regardless of the value of the penalty parameter.

6.2 The functional to be minimized

From the geometrical standpoint, the interface is treated as a union of coinciding faces of the general hexahedral finite elements sharing the hanging nodes as shown in Figure 6.1. One of the faces is refered to as the *slave surface* while the other face is refered to as the *master surface.* Let us refer to elements containing the slave surfaces as *slave elements* and elements containing the master surfaces as *master elements.* Let $\mathbf{x}_s$ represent a generic point on the surface of slave element $\mathscr{E}_s$ containing the slave surface and let $\mathbf{x}_m$ represent the *orthogonal projection* of $\mathbf{x}_s$ onto the surface of master element $\mathscr{E}_m$ containing the master surface. Let $\mathbf{u}_s$ and $\mathbf{u}_m$ represent the displacement field evaluated at $\mathbf{x}_s$ and $\mathbf{x}_m$ respectively. Let $\mathscr{C}^s$ and $\mathscr{C}^m$ be the strain energies of $\mathscr{E}_s$ and $\mathscr{E}_m$ respectively. Then the augmented functional to be minimized is

$$\tilde{\mathscr{C}} \equiv \overbrace{\mathscr{C}^s + \mathscr{C}^m}^{\textcircled{1}} + \overbrace{\int\limits_{\Gamma_c} \boldsymbol{\lambda} \cdot \mathbf{g}\, dA}^{\textcircled{2}} + \overbrace{\frac{1}{2} \int\limits_{\Gamma_c} \epsilon \mathbf{g} \cdot \mathbf{g}\, dA}^{\textcircled{3}}$$

where $\mathbf{g} \equiv \mathbf{u}_s - \mathbf{u}_m$ is the refered to as the *penetration function*, $\boldsymbol{\lambda} \equiv \frac{1}{2}(\mathbf{t}^s + \mathbf{t}^m)$ is the force conjugate to the constraint $\mathbf{g} = \mathbf{0}$; introduced in a mean sense, and $\mathbf{t}^s$ and $\mathbf{t}^m$ are force conjugates to $\mathbf{g}$ at $\mathbf{x}_s$ and $\mathbf{x}_m$ respectively. The term ① is the total strain energy of $\mathscr{E}_s$ and $\mathscr{E}_m$, the term ② is the lagrange multiplier term and the term ③ is the penalty term. The term ② enforces the constraint $\mathbf{g} = \mathbf{0}$ via lagrange multipliers, and the term ③ penalizes any deviation from the constraint $\mathbf{g} = \mathbf{0}$. The integrals are evaluated with respect to one of the surfaces (in this case surface of $\mathscr{E}_s$ or the slave surface). *In essence, the slave elements are all the fine mesh finite elements sharing the interface.*

6.2.1 Orthogonal projections

Let $\mathbf{X}_{m_i}$, $i = 1, .., 8$ and $\mathbf{X}_{s_j}$, $j = 1, .., 8$ be the coordinates of the finite element nodes of $\mathscr{E}_s$ and $\mathscr{E}_m$ respectively. Let (ξ, η, μ) represent the spatial field in reference element $\hat{\mathscr{E}}$ and let $N_i(\xi, \eta, \mu)$, $i = 1, .., 8$ represent the shape functions. Then we have

$$\left\{ \begin{array}{c} \mathbf{x}_m = \sum\limits_{i=1}^{8} N_i|_{(\xi_m,\eta_m,\mu_m)} \mathbf{X}_{m_i} \\ \mathbf{x}_s = \sum\limits_{i=1}^{8} N_i|_{(\xi_m,\eta_m,\mu_m)} \mathbf{X}_{s_i} \end{array} \right\} \tag{6.1}$$

where (ξ_m, η_m, μ_m) and (ξ_s, η_s, μ_s) are the coordinates of $\mathbf{x}_m$ and $\mathbf{x}_s$ respectively mapped onto the reference element $\hat{\mathscr{E}}$. It is critical to note that (ξ_s, η_s, μ_s) is known while (ξ_m, η_m, μ_m) is to be determined. The components $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$ of the tangent at $\mathbf{x}_m$ with respect to the local axis of master surface are computed as

$$\left\{ \begin{array}{c} \mathbf{e}_1 = \sum\limits_{i=1}^{8} \frac{\partial N_i}{\partial \xi}|_{(\xi_m,\eta_m,\mu_m)} \mathbf{X}_{m_i} \\ \mathbf{e}_2 = \sum\limits_{i=1}^{8} \frac{\partial N_i}{\partial \eta}|_{(\xi_m,\eta_m,\mu_m)} \mathbf{X}_{m_i} \\ \mathbf{e}_3 = \sum\limits_{i=1}^{8} \frac{\partial N_i}{\partial \mu}|_{(\xi_m,\eta_m,\mu_m)} \mathbf{X}_{m_i} \end{array} \right\} \tag{6.2}$$

$\mathbf{x}_m$ is the orthogonal projection onto the master surface of $\mathbf{x}_s$ on the slave surface. The orthogonality condition is satisfied by

$$\left\{ \begin{array}{c} \mathbf{e}_1 \cdot (\mathbf{x}_s - \mathbf{x}_m) = 0 \\ \mathbf{e}_2 \cdot (\mathbf{x}_s - \mathbf{x}_m) = 0 \\ \mathbf{e}_3 \cdot (\mathbf{x}_s - \mathbf{x}_m) = 0 \end{array} \right\} \tag{6.3}$$

Substituting (6.1) and (6.2) in (6.3), we get

$$\left\{ \begin{array}{l} \sum\limits_{j=1}^{8} \frac{\partial N_j}{\partial \xi}\big|_{(\xi_m,\eta_m,\mu_m)} \mathbf{X}_{m_j} \cdot \left(\mathbf{x}_s - \sum\limits_{k=1}^{8} N_k\big|_{(\xi_m,\eta_m,\mu_m)} \mathbf{X}_{m_k} \right) \equiv f_1(\xi_m, \eta_m, \mu_m) = 0 \\ \sum\limits_{j=1}^{8} \frac{\partial N_j}{\partial \eta}\big|_{(\xi_m,\eta_m,\mu_m)} \mathbf{X}_{m_j} \cdot \left(\mathbf{x}_s - \sum\limits_{k=1}^{8} N_k\big|_{(\xi_m,\eta_m,\mu_m)} \cdot \mathbf{X}_{m_k} \right) \equiv f_2(\xi_m, \eta_m, \mu_m) = 0 \\ \sum\limits_{j=1}^{8} \frac{\partial N_j}{\partial \mu}\big|_{(\xi_m,\eta_m,\mu_m)} \mathbf{X}_{m_j} \cdot \left(\mathbf{x}_s - \sum\limits_{k=1}^{8} N_k\big|_{(\xi_m,\eta_m,\mu_m)} \cdot \mathbf{X}_{m_k} \right) \equiv f_2(\xi_m, \eta_m, \mu_m) = 0 \end{array} \right\} \tag{6.4}$$

The solution to (6.4) is obtained iteratively for the $(k+1)^{th}$ iteration as

$$\left\{ \begin{array}{c} \xi_m \\ \eta_m \\ \mu_m \end{array} \right\}^{k+1} = \left\{ \begin{array}{c} \xi_m \\ \eta_m \\ \mu_m \end{array} \right\}^{k} - \begin{bmatrix} \frac{\partial f_1}{\partial \xi} & \frac{\partial f_1}{\partial \eta} & \frac{\partial f_1}{\partial \mu} \\ \frac{\partial f_2}{\partial \xi} & \frac{\partial f_2}{\partial \eta} & \frac{\partial f_2}{\partial \mu} \\ \frac{\partial f_3}{\partial \xi} & \frac{\partial f_3}{\partial \eta} & \frac{\partial f_3}{\partial \mu} \end{bmatrix}^{-1} \left\{ \begin{array}{c} f_1 \\ f_2 \\ f_3 \end{array} \right\}$$

with the initial guess as $\left\{ \begin{array}{c} \xi_m \\ \eta_m \\ \mu_m \end{array} \right\}^{0} = \left\{ \begin{array}{c} 0 \\ 0 \\ 0 \end{array} \right\}$ and the RHS being evaluated based on the values $\left\{ \begin{array}{c} \xi_m \\ \eta_m \\ \mu_m \end{array} \right\}^{k}$ obtained from the previous k^{th} iteration. The stopping criterion is

$$\| \left\{ \begin{array}{c} \xi_m \\ \eta_m \\ \mu_m \end{array} \right\}^{k+1} - \left\{ \begin{array}{c} \xi_m \\ \eta_m \\ \mu_m \end{array} \right\}^{k} \| < TOL * \| \left\{ \begin{array}{c} \xi_m \\ \eta_m \\ \mu_m \end{array} \right\}^{k} \|$$

where TOL is a pre-specified tolerance.

6.3 Variation of the functional

Let $\mathbf{U}$ be the vector of displacement degrees of freedom at the finite element nodes and let $\mathbf{P}$ represent the vector of nodal forces. Then the system of equations after the minimization of the augmented energy functional would be

$$\mathbf{K}\mathbf{U} = \mathbf{P} \tag{6.5}$$

where $\mathbf{K}$ is the stiffness matrix. We rearrange the vector $\mathbf{U}$ in the following form

$$\mathbf{U} = \left\{ \begin{array}{c} \mathbf{U}_r \\ \mathbf{U}_s \\ \mathbf{U}_m \end{array} \right\}$$

where $\mathbf{U}_s$ are the displacement degrees of freedom corresponding to all the slave element nodes, $\mathbf{U}_m$ are the displacement degrees of freedom corresponding to all the master element nodes and $\mathbf{U}_r$ are the displacement degrees of freedom corresponding to all the remaining nodes on the finite element mesh. The system of equations (6.5) is then written as

$$\left[\mathbf{K}^d + \begin{bmatrix} . & . & . \\ . & \mathbf{K}_{ss} & \mathbf{K}_{sm} \\ . & \mathbf{K}_{ms} & \mathbf{K}_{mm} \end{bmatrix} \right] \left\{ \begin{array}{c} \mathbf{U}_r \\ \mathbf{U}_s \\ \mathbf{U}_m \end{array} \right\} = \mathbf{P} \tag{6.6}$$

where $\mathbf{K}^d$ is the stiffness matrix that is obtained after the minimization of the original energy functional. The objective is to obtain expressions for the submatrices $\mathbf{K}_{ss}$, $\mathbf{K}_{sm}$, $\mathbf{K}_{ms}$ and $\mathbf{K}_{mm}$ that arise as a result of the minimization of additional terms (lagrange multiplier term and penalty term) in the augmented energy functional. For the sake of clarity, we rewrite the augmented Lagrangian functional as follows

$$\tilde{\mathscr{C}} \equiv \mathscr{C}^s + \mathscr{C}^m + \int\limits_{\Gamma_c} \frac{1}{2}(\mathbf{t}^s + \mathbf{t}^m) \cdot \mathbf{g}\, dA + \frac{1}{2} \int\limits_{\Gamma_c} \epsilon \mathbf{g} \cdot \mathbf{g}\, dA$$

The first variation of the energy functional would be

$$\delta\tilde{\mathscr{C}} = \delta(\mathscr{C}^s + \mathscr{C}^m) \\ + \overbrace{\int\limits_{\Gamma_c} \frac{1}{2}(\delta\mathbf{t}^s + \delta\mathbf{t}^m)\cdot\mathbf{g}\,dA + \int\limits_{\Gamma_c} \frac{1}{2}(\mathbf{t}^s + \mathbf{t}^m)\cdot\delta\mathbf{g}\,dA + \int\limits_{\Gamma_c} \epsilon\mathbf{g}\cdot\delta\mathbf{g}\,dA}^{\mathcal{C}}$$

which can also be written as

$$\delta\tilde{\mathscr{C}} = \delta\mathbf{U}^T(\mathbf{K}^d\mathbf{U} - \mathbf{P}) \\ + \overbrace{\int\limits_{\Gamma_c} \frac{1}{2}(\delta\mathbf{t}^s + \delta\mathbf{t}^m)\cdot\mathbf{g}\,dA + \int\limits_{\Gamma_c} \frac{1}{2}(\mathbf{t}^s + \mathbf{t}^m)\cdot\delta\mathbf{g}\,dA + \int\limits_{\Gamma_c} \epsilon\mathbf{g}\cdot\delta\mathbf{g}\,dA}^{\mathcal{C}} \tag{6.7}$$

6.4 Numerical integration for evaluation of surface integrals

Let $\mathcal{E}_s$ respresent the set of slave elements $\mathscr{E}_s$. The contribution to the integral $\mathcal{C}$ in Equation (6.7) over every slave surface is evaluated as a sum of the integrands evaluated at the four gauss points multiplied by the determinant of the jacobian of the mapping from the reference 2D element to the slave surface as follows

$$\mathcal{C} = \sum_{\mathscr{E}_s \in \mathcal{E}_s} \left[\sum_{N=1}^{4} \left[\frac{1}{2}(\delta\mathbf{t}^s + \delta\mathbf{t}^m)\cdot\mathbf{g} + \frac{1}{2}(\mathbf{t}^s + \mathbf{t}^m)\cdot\delta\mathbf{g} + \epsilon\mathbf{g}\cdot\delta\mathbf{g} \right] detJ_{\mathscr{E}_s} \right] \tag{6.8}$$

where $|\mathcal{E}_s|$ is the number of slave surfaces, and is equal to the number of slave elements. *The determinant of the jacobian varies from slave surface to slave surface since every slave surface belongs to a different slave element, and hence the jacobian of the mapping from the reference 2D element to the slave element* *varies from one element to the other.* Now, corresponding to each gauss point on the reference element to which the slave surface is mapped onto, there is an actual physical point on the slave surface. Let's refer to that point as $\mathbf{x}_s$. We employ the logic elucidated in module 6.2.1 to evaluate the orthogonal projection $\mathbf{x}_m$ of $\mathbf{x}_s$ onto the master surface.

6.4.1 Evaluating the force conjugates one gauss point at a time

The force conjugate to the constraint evaluated at (ξ_s, η_s, μ_s) is given by

$$\begin{aligned}
\mathbf{t}^s &= \begin{bmatrix} \sigma_1 & \sigma_4 & \sigma_6 \\ \sigma_4 & \sigma_2 & \sigma_5 \\ \sigma_6 & \sigma_5 & \sigma_3 \end{bmatrix} \Bigg|_{(\xi_s,\eta_s,\mu_s)} \begin{Bmatrix} n_1 \\ n_2 \\ n_3 \end{Bmatrix} \Bigg|_{(\xi_s,\eta_s,\mu_s)} \\
&\equiv \begin{bmatrix} n_1 & 0 & 0 & n_2 & 0 & n_3 \\ 0 & n_2 & 0 & n_1 & n_3 & 0 \\ 0 & 0 & n_3 & 0 & n_2 & n_1 \end{bmatrix} \Bigg|_{(\xi_s,\eta_s,\mu_s)} \begin{Bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \sigma_4 \\ \sigma_5 \\ \sigma_6 \end{Bmatrix} \Bigg|_{(\xi_s,\eta_s,\mu_s)} \\
&\equiv \overbrace{\begin{bmatrix} n_1 & 0 & 0 & n_2 & 0 & n_3 \\ 0 & n_2 & 0 & n_1 & n_3 & 0 \\ 0 & 0 & n_3 & 0 & n_2 & n_1 \end{bmatrix} \Bigg|_{(\xi_s,\eta_s,\mu_s)} \mathcal{D}\mathbf{B}|_{(\xi_s,\eta_s,\mu_s)}}^{\mathcal{F}_{\mathscr{E}_s}} \mathbf{U}_{s|_{\mathscr{E}_s}} \qquad (6.9)
\end{aligned}$$

where $\begin{Bmatrix} n_1 \\ n_2 \\ n_3 \end{Bmatrix} \Bigg|_{(\xi_s,\eta_s,\mu_s)}$ is the normal to the slave surface evaluated at (ξ_s, η_s, μ_s), $\mathcal{D}$ is the 6×6 constitutive matrix, $\mathbf{B}|_{(\xi_s,\eta_s,\mu_s)}$ is the 6×24 strain displacement interpolation matrix evaluated at (ξ_s, η_s, μ_s) and $\mathbf{U}_{s|_{\mathscr{E}_s}}$ is the restriction of $\mathbf{U}_s$ to slave element $\mathscr{E}_s$. Similarly, the force conjugate to the constraint evaluated at (ξ_m, η_m, μ_m) is given by

$$\mathbf{t}^m \equiv \overbrace{\left.\begin{bmatrix} n_1 & 0 & 0 & n_2 & 0 & n_3 \\ 0 & n_2 & 0 & n_1 & n_3 & 0 \\ 0 & 0 & n_3 & 0 & n_2 & n_1 \end{bmatrix}\right|_{(\xi_m,\eta_m,\mu_m)} \mathcal{D}\mathbf{B}|_{(\xi_m,\eta_m,\mu_m)}}^{\mathcal{F}_{\mathscr{E}_m}} \mathbf{U}_{m|_{\mathscr{E}_m}} \tag{6.10}$$

where $\left.\begin{Bmatrix} n_1 \\ n_2 \\ n_3 \end{Bmatrix}\right|_{(\xi_m,\eta_m,\mu_m)}$ is the normal to the master surface evaluated at (ξ_m, η_m, μ_m), $\mathbf{B}|_{(\xi_m,\eta_m,\mu_m)}$ is the 6×24 strain displacement interpolation matrix evaluated at (ξ_m, η_m, μ_m) and and $\mathbf{U}_{m|_{\mathscr{E}_m}}$ is the restriction of $\mathbf{U}_m$ to master element $\mathscr{E}_m$. The normals $\left.\begin{Bmatrix} n_1 \\ n_2 \\ n_3 \end{Bmatrix}\right|_{(\xi_s,\eta_s,\mu_s)}$ and $\left.\begin{Bmatrix} n_1 \\ n_2 \\ n_3 \end{Bmatrix}\right|_{(\xi_m,\eta_m,\mu_m)}$ are obtained as follows

$$\begin{aligned} \left.\begin{Bmatrix} n_1 \\ n_2 \\ n_3 \end{Bmatrix}\right|_{(\xi_s,\eta_s,\mu_s)} &= \left.\frac{\nabla S_s}{\|\nabla S_s\|}\right|_{(\xi_s,\eta_s,\mu_s)} \\ \left.\begin{Bmatrix} n_1 \\ n_2 \\ n_3 \end{Bmatrix}\right|_{(\xi_m,\eta_m,\mu_m)} &= \left.\frac{\nabla S_m}{\|\nabla S_m\|}\right|_{(\xi_m,\eta_m,\mu_m)} \end{aligned}$$

where S_s and S_m are equations of the slave and master surfaces respectively. The equations of the surfaces given coordinates of the four points are obtained using the procedure of singular value decompositions as described in Module 3.1.3.1.

6.4.2 Evaluating the penetration function one gauss point at a time

The penetration function is given by

$$\mathbf{g} = \mathbf{u}_s - \mathbf{u}_m = \overbrace{\mathbf{N}|_{(\xi_s,\eta_s,\mu_s)}}^{\mathcal{N}_{\mathscr{E}_s}} \mathbf{U}_{s|_{\mathscr{E}_s}} - \overbrace{\mathbf{N}|_{(\xi_m,\eta_m,\mu_m)}}^{\mathcal{N}_{\mathscr{E}_m}} \mathbf{U}_{m|_{\mathscr{E}_m}} \tag{6.11}$$

where $\mathbf{N}$ is the 3×24 shape function matrix

6.4.3 Evaluating the surface integral

In lieu of Equations (6.9) - (6.11), the surface integral (6.8) is evaluated as

$$
\begin{aligned}
\mathcal{C} = \sum_{\mathscr{E}_s \in \mathcal{E}_s} \Bigg[\sum_{N=1}^{4} \Bigg[& \frac{1}{2}(\mathcal{F}_{\mathscr{E}_s}\delta\mathbf{U}_{s|\mathscr{E}_s} + \mathcal{F}_{\mathscr{E}_m}\delta\mathbf{U}_{m|\mathscr{E}_m}) \cdot (\mathcal{N}_{\mathscr{E}_s}\mathbf{U}_{s|\mathscr{E}_s} - \mathcal{N}_{\mathscr{E}_m}\mathbf{U}_{m|\mathscr{E}_m}) \\
& + \frac{1}{2}(\mathcal{F}_{\mathscr{E}_s}\mathbf{U}_{s|\mathscr{E}_s} + \mathcal{F}_{\mathscr{E}_m}\mathbf{U}_{m|\mathscr{E}_m}) \cdot (\mathcal{N}_{\mathscr{E}_s}\delta\mathbf{U}_{s|\mathscr{E}_s} - \mathcal{N}_{\mathscr{E}_m}\delta\mathbf{U}_{m|\mathscr{E}_m}) \\
& + \epsilon(\mathcal{N}_{\mathscr{E}_s}\mathbf{U}_{s|\mathscr{E}_s} - \mathcal{N}_{\mathscr{E}_m}\mathbf{U}_{m|\mathscr{E}_m}) \cdot (\mathcal{N}_{\mathscr{E}_s}\delta\mathbf{U}_{s|\mathscr{E}_s} - \mathcal{N}_{\mathscr{E}_m}\delta\mathbf{U}_{m|\mathscr{E}_m}) \Bigg] detJ_{\mathscr{E}_s} \Bigg]
\end{aligned}
$$

which can also be written as

$$
\begin{aligned}
\mathcal{C} = {} & \delta\mathbf{U}_s^T \overbrace{\sum_{\mathscr{E}_s \in \mathcal{E}_s} \Bigg[\overbrace{\sum_{N=1}^{4} \Bigg[\frac{1}{2}\mathcal{F}_{\mathscr{E}_s}^T \mathcal{N}_{\mathscr{E}_s} + \frac{1}{2}\mathcal{N}_{\mathscr{E}_s}^T \mathcal{F}_{\mathscr{E}_s} + \epsilon \mathcal{N}_{\mathscr{E}_s}^T \mathcal{N}_{\mathscr{E}_s} \Bigg] detJ_{\mathscr{E}_s}}^{\mathbf{K}_{ss|\mathscr{E}_s}} \Bigg]}^{\mathbf{K}_{ss}} \mathbf{U}_s \\
& + \delta\mathbf{U}_s^T \overbrace{\sum_{\mathscr{E}_s \in \mathcal{E}_s} \Bigg[\overbrace{\sum_{N=1}^{4} \Bigg[-\frac{1}{2}\mathcal{F}_{\mathscr{E}_s}^T \mathcal{N}_{\mathscr{E}_m} + \frac{1}{2}\mathcal{N}_{\mathscr{E}_s}^T \mathcal{F}_{\mathscr{E}_m} - \epsilon \mathcal{N}_{\mathscr{E}_s}^T \mathcal{N}_{\mathscr{E}_m} \Bigg] detJ_{\mathscr{E}_s}}^{\mathbf{K}_{sm|\mathscr{E}_s}} \Bigg]}^{\mathbf{K}_{sm}} \mathbf{U}_m \\
& + \delta\mathbf{U}_m^T \overbrace{\sum_{\mathscr{E}_s \in \mathcal{E}_s} \Bigg[\overbrace{\sum_{N=1}^{4} \Bigg[\frac{1}{2}\mathcal{F}_{\mathscr{E}_m}^T \mathcal{N}_{\mathscr{E}_s} - \frac{1}{2}\mathcal{N}_{\mathscr{E}_m}^T \mathcal{F}_{\mathscr{E}_s} - \epsilon \mathcal{N}_{\mathscr{E}_m}^T \mathcal{N}_{\mathscr{E}_s} \Bigg] detJ_{\mathscr{E}_s}}^{\mathbf{K}_{ms|\mathscr{E}_s}} \Bigg]}^{\mathbf{K}_{ms}} \mathbf{U}_s \\
& + \delta\mathbf{U}_m^T \overbrace{\sum_{\mathscr{E}_s \in \mathcal{E}_s} \Bigg[\overbrace{\sum_{N=1}^{4} \Bigg[-\frac{1}{2}\mathcal{F}_{\mathscr{E}_m}^T \mathcal{N}_{\mathscr{E}_m} - \frac{1}{2}\mathcal{N}_{\mathscr{E}_m}^T \mathcal{F}_{\mathscr{E}_m} + \epsilon \mathcal{N}_{\mathscr{E}_m}^T \mathcal{N}_{\mathscr{E}_m} \Bigg] detJ_{\mathscr{E}_s}}^{\mathbf{K}_{mm|\mathscr{E}_s}} \Bigg]}^{\mathbf{K}_{mm}} \mathbf{U}_m
\end{aligned}
\tag{6.12}
$$

where $\mathbf{K}_{ss}$, $\mathbf{K}_{sm}$, $\mathbf{K}_{ms}$ and $\mathbf{K}_{mm}$ are obtained after assembling the contributions $\mathbf{K}_{mm|_{\mathscr{E}_s}}$, $\mathbf{K}_{sm|_{\mathscr{E}_s}}$, $\mathbf{K}_{ms|_{\mathscr{E}_s}}$ and $\mathbf{K}_{ss|_{\mathscr{E}_s}}$ from each slave element $\mathscr{E}_s$.

6.5 System of Equations

The system of equations is obtained by equating the variation of the functional to zero as follows

$$\begin{aligned} \delta\tilde{\mathscr{C}} \equiv \delta\mathbf{U}^T(\mathbf{K}^d\mathbf{U} - \mathbf{P}) + \delta\mathbf{U_s}^T\mathbf{K}_{ss}\mathbf{U}_s + \delta\mathbf{U_s}^T\mathbf{K}_{sm}\mathbf{U}_m \\ + \delta\mathbf{U_m}^T\mathbf{K}_{ms}\mathbf{U}_s + \delta\mathbf{U_m}^T\mathbf{K}_{mm}\mathbf{U}_m = 0 \end{aligned}$$

which is eventually written as

$$\left[\mathbf{K}^d + \begin{bmatrix} . & . & . \\ . & \mathbf{K}_{ss} & \mathbf{K}_{sm} \\ . & \mathbf{K}_{ms} & \mathbf{K}_{mm} \end{bmatrix}\right] \begin{Bmatrix} \mathbf{U}_r \\ \mathbf{U}_s \\ \mathbf{U}_m \end{Bmatrix} = \mathbf{P}$$

where $\mathbf{K}_{ss}$, $\mathbf{K}_{sm}$, $\mathbf{K}_{ms}$ and $\mathbf{K}_{mm}$ are given in Equation (6.12).

6.6 Procedural framework

The steps to be followed for the treatment of hanging nodes in hexahedral meshes are

- Identify the elements sharing the interface
- Identify the elements on the fine mesh side as slave elements and elements on the coarse mesh side as master elements
- Identify the faces of the slave elements on the interface as slave surfaces and faces of the master elements on the interface as master surfaces

- Use singular value decompositions (see module 3.1.3.1) to obtain the equations of the slave and master surfaces
- In the numerical integration module, map the slave and master surfaces to 2D reference elements
- For every gauss point on the reference element which every slave surface has been mapped onto, identify the point on the slave surface. Use the equation of the slave surface to obtain the normal to the slave surface at that point.
- Obtain the orthogonal projection of that point onto the master surface. Use the equation of the master surface to obtain the normal to the master surface at that point.
- Obtain the contributions to the submatrices from each slave element
- Assemble the contributions to obtain the global submatrices

6.7 Summary

We present an augmented Lagrangian formulation for treatment of hanging nodes in hexahedral meshes. The motivation for this work lies in the quest in addressing some of the limitations of the two grid staggered solution algorithmic framework put in place in Chapters 2 and 3. The augmented Lagrangian is one of the techniques employed in the field of computational contact mechanics

Chapter 7

Conclusions and scope for future work

7.1 Conclusions

The two-grid algorithm we developed was a consequence of the need to resolve the differing length scales of the different physics as accurately as possible, while keeping the problem computationally feasible. The algorithm was developed as an addendum to the existing staggered solution procedure that solves the different physics on the same length scales sequentially and iteratively until convergence. We demostrated that the problem of incorporating overburden in a reservoir simulator can be solved, albeit with recourse to concepts in computational geometry and parallel computing. There are indeed other approaches, among the many possibilities, of solving the problem, and it would be interesting to see how much the physics of the overburden plays a role in phenomena inside the reservoir. Fractures that do not propagate, as well as fractures that do propagate would be the logical next step as far as a problem statement to be tackled is concerned. The problem with fractures lies in the fact that they are lower dimensional features compared to the surrounding problem set-up (for example, a 1D fracture in a 2D problem). The augmented Lagrangian we propose in Chapter 6 does handle hanging nodes, but it only solves the scalability problem in incorporating fractures in a big domain (implying that in the absence of hanging nodes, the problem of treating fractures in a realistic subsurface framework might be impossibly expensive). There are obviously considerations of the finite element discretization employed that yields sufficiently accurate results: the mixed finite element method for flow yields local mass conservation (which means that the fluid mass does not vanish when we traverse one finite element to the other); and we employ conforming finite element method for the mechanics which does not preserve momentum element-by-element. There are techniques to force the momentum to be preserved element-by-element: one is to use a mixed finite element method for the mechanics problem as well, the other is to move towards a finite element method that allows for discontinuities in the primary variable, thereby allowing for enough leeway in preserving momentum locally. All-in-all, the geophysics-reservoir problem is a fine balance among incorporating all much physics as possible, preserving the physics that has been incorporated, and not blowing up the supercomputer in the process.

7.2 Scope for future work

7.2.1 Fixed stress split algorithm

We have already commented in module 4.3.1.1 that the fixed stress split algorithm provides an inconsistency during the flow solve in the first coupling iteration in every time step. The authors have not ventured out to resolve that inconsistency yet as the algorithm performs smoothly and is accurate for the case of slightly compressible single phase flow coupled with linear poromechanics. The possible premise is that the inconsistency does not hamper the solution accuracy as long as the number of fixed stress iterations per time step is sufficient for the solution pollution to die out. That might not be the case when the nonlinearities in the flow model are no longer mild as they are in this dissertation. With that in mind, there is scope for modification of the logic involved in the first coupling iteration, or indeed modification of the staggered solution framework itself, for the case when the flow model is more sophisticated

7.2.2 Link between staggering and homogenization

The link we established in the latter half of Chapter 4 between the measure of the mean stress used in the decoupling constraint for the two-grid algorithm and the Voigt and Reuss bounds has interesting connotations for the imposed homogeneous boundary conditions used to arrive at effective properties in the computational homogenization of multiphase composites. We know that stress uniform boundary conditions on the mesoscale lead to the Reuss bound on the effective property at the macroscale while the kinematic uniform boundary conditions on the mesoscale lead to the Voigt bound on the effective property at the macroscale (Hashin and Shtrikman [44], Hill [45], Hill [46], Hill [47], Hashin [43], Zohdi and Wriggers [102]). We also know that periodic boundary conditions on the mesoscale lead to the most accurate effective properties at the macroscale. In case of the two-grid approach, the fine scale flow grid is the mesoscale while the coarse scale poromechanical grid is the macroscale. When the adjustable parameter takes upon the value of twice the drained bulk modulus, we obtain the Reuss bound corresponding to stress uniform boundary conditions on the mesoscale. Similarly, when the adjustable parameter takes upon the value of the inverse of the drained bulk modulus, we obtain the Voigt bound corresponding to kinematic uniform boundary conditions on the mesoscale. By an extension of that logic, we expect a certain value of the adjustable parameter that corresponds to the periodic boundary conditions imposed on the mesoscale thereby lending itself to the most accurate estimate of the macroscale effective property. At the same time, when the value of the adjustable paramter is twice the bulk modulus, we observe the fastest convergence of the algorithm. This implies that the optimal choice of the adjustable parameter depends on the compromise between accuracy and speed. It is yet to be hypothesized what that choice might be.

7.2.3 Treatment of rock anisotropy

In Chapter 5, we presented a staggered solution algorithm that incorporates rock anisotropy, albeit with flow and mechanics being resolved on the same grid. It would be interesting to see how the algorithm pans out with the two-grid approach.

7.2.4 Treatment of hanging nodes

We also commented at the start of Chapter 6 that the algorithm we developed in Chapters 2 and 3 in restrictive for problems that have fine scale features in the poromechanics problem. With that in mind, we formulated an augmented Lagrangian that treats hanging nodes in hexahedral meshes with an intent to use that in the poromechanics finite element module. For the lack of time, the augmented Lagrangian has not been implemented yet in IPARS, and it would be interesting to see how that code framework addendum would affect memory management and how it would solve problems involving fine scale features for geomechanics in addition to providing uplift and subsidence calculations, as well as mimicking the true effect of overburden and sideburdens on the reservoir.

7.2.5 Extension to poroelastoplasticity

This dissertation deals with slightly compressible single flow coupled with linear poromechanics for both isotropic and anisotropic porous media, and does not venture out to resolve possible issues pertaining to the use of the fixed stress split strategy for the solution of the slightly compressible single phase coupled with nonlinear poromechanics. Although we do make a start in that direction by extending the convergence analysis framework previously employed in Chapters 4 and 5 for this problem too

- S. Dana and M. F. Wheeler, ‘Convergence analysis of fixed stress split iterative scheme for small strain anisotropic poroelastoplasticity: a primer”, *arXiv* 1810.04163

this analysis is incomplete unless certain mathematical arguments are invoked. A list of suggestions to complete the analysis are provided in the document.

Indeed, it is possible that the fixed stress split strategy may fall apart for this case, which would again suggest potential improvements to the solution logic when nonlinear poromechanics is invoked

7.2.6 Locally momentum conserving discretization for mechanics

We employed continuous Galerkin finite element method for the poromechanics subproblem in this dissertation. This approach lends to locally non-conservative momenta, implying that there is loss in momentum while traversing from one finite element to the other. One possible solution is to employ discontinuous Galerkin finite element method or a degenerate enriched Galerkin method.

7.2.7 Invoking different time steps for flow and poromechanics

In this dissertation, we exploited the differing length scales of flow and geomechanics to come up with an algorithm that solves the two physics on different finite element grids. It would be interesting to see how it combines with the idea of exploiting the differing time scales of the two physics using different time stepping procedures.

Appendix A

A historical review of theory of deformable porous media, finite element modeling and solution algorithms

The discovery of fundamental mechanical effects in saturated deformable porous solids, and the formulation of the first porous media theories, are mainly due to two distinguished professors at the Technische Hochschule of Vienna, namely Paul Fillunger and Karl von Terzaghi. Fillunger [33] was the first to state that the pore liquid pressure does not have any influence upon the strength of the porous solid. Besides this comment, he remarked that the pore water pressure does not affect the material behaviour of the porous solid at all. Terzaghi [84] started the development of the idea of 'effective stress' within the framework of the treatment of the consolidation problem for clay layers. According to Skempton [80], the formalism of the idea of effective stress was finally given by Terzaghi [85]. The basic idea of the principle of effective stress is that the 'effective' stress responsible for causing soil deformation is the excess of the total stress over the pore fluid pressure. This implies that the linear momentum balance for poroelasticity can be obtained by simply replacing the total stress in the linear momentum balance for elasticity by the effective stress. In 1941, a Belgian physicist named Maurice Anthony Biot extended Terzaghi's one - dimensional theory to the three - dimensional case and further introduced a parameter representative of the degree of saturation of the fluid inside the pores of the solid (see Biot [8]). He assumed the following properties of soils

- isotropy of the material,
- reversibility of stress-strain relations under final equilibrium conditions,
- linearity of stress-strain relations,
- small strains,
- the water contained in the pores is incompressible,
- the water may contain air bubbles,
- the water flows through the porous skeleton according to Darcy's law

With E, G, ν representing the Young's modulus, shear modulus, and Poisson's ratio respectively, θ the increment of water volume per unit volume of soil, σ the increment of water pressure, $\mathbf{T}^S$ the stress, and $\mathbf{u}$ the soil displacement, he introduced coefficients H, R and α and arrived at the following relations

$$\mathbf{T}^S = 2G\left(\mathbf{E}^L + \frac{\nu}{1-2\nu}(\operatorname{div}\mathbf{u})\,\mathbf{I}\right) - \alpha\sigma\,\mathbf{I} \tag{A.1}$$

$$\theta = \alpha\operatorname{div}\mathbf{u} + \sigma\left(\frac{1}{R} - \frac{\alpha}{H}\right) \tag{A.2}$$

where $\mathbf{E}^L$ is the well-known linearized Green strain tensor given by

$$\mathbf{E}^L = \frac{1}{2}(\operatorname{grad}\mathbf{u} + (\operatorname{grad}\mathbf{u})^T)$$

The corresponding stresses must satisfy the equilibrium conditions of a stress field (inertia forces and body forces are neglected):

$$\operatorname{div} \mathbf{T}^S = \mathbf{0} \tag{A.3}$$

The relation (A.1) along with the conditions (A.3) result in

$$G \operatorname{div} \operatorname{grad} \mathbf{u} + (\lambda + G) \operatorname{grad} \operatorname{div} \mathbf{u} - \alpha \operatorname{grad} \sigma = \mathbf{0} \tag{A.4}$$

where $\lambda \equiv \frac{E\nu}{(1+\nu)(1-2\nu)}$ is a Lamé parameter. In order to have a complete set of equations for determining $\mathbf{u}$ and σ, one more equation is necessary. This equation can be obtained by introducing Darcy's law[1] governing the flow of water in a porous medium:

$$\mathbf{w}^F = -k^F \operatorname{grad} \sigma \tag{A.5}$$

where $\mathbf{w}^F$ represents the volume of water flowing per second and unit area through the faces of an elementary cube and k^F denotes the permeability coefficient of the soil. Furthermore, under the assumption of the incompressibility of water, the rate of water content of an element of soil must be equal to the volume of water entering per second through the surface of the elements.

$$\frac{\partial \theta}{\partial t} = -\operatorname{div} \mathbf{w}^F \tag{A.6}$$

---

[1] Darcy [27] observed in tests with natural sand, the proportionality of the total volume of water running through the sand and the loss of pressure. The relationship he arrived at is called the 'Darcy's law' and is repeatedly used as the equation that governs the flow of fluid in porous media.

Combining (A.2), (A.5) and (A.6),

$$k^F \operatorname{div} \operatorname{grad} \sigma = \alpha \frac{\partial(\mathbf{E}^L \cdot \mathbf{I})}{\partial t} + \left(\frac{1}{R} - \frac{\alpha}{H} \right) \frac{\partial \sigma}{\partial t} \tag{A.7}$$

(A.4) and (A.7) are the basic relations of Biot's theory of consolidation, where (A.4) describes the settlement of the soil, and (A.7) describes the change in the water pressure. Furthermore, it is important to note that these two relations are coupled. Biot [9] relaxed the assumption of incompressibility of the pore fluid and isotropy of the pore skeleton in order to render a set of equations applicable to an anisotropic porous solid containing a viscous compressible fluid. The resulting set of equations for the specific case of isotropic porous solid had the same number (three) of coefficients to be determined as the original equations (A.4) and (A.7). Biot and Willis [10] reworked the set of equations with a reduced number (two) of coefficients and further devised tests[2] to determine those coefficients. The authors summarized the equations for the isotropic case (inertia forces and body forces are neglected):

$$\operatorname{div} \left(2G\mathbf{E}^L + \lambda \left(\operatorname{div} \mathbf{u} \right) \mathbf{I} - \alpha p \mathbf{I} \right) = \mathbf{0} \tag{A.8}$$

$$\frac{\partial \zeta}{\partial t} + \operatorname{div} \mathbf{w}^F = 0 \tag{A.9}$$

$$\mathbf{w}^F = -k^F \operatorname{grad} p \tag{A.10}$$

---

[2] The interested reader is refered to the review article of Detournay and Cheng [28] for a brief description of these tests.

where ζ and $\boldsymbol{\sigma}$ are the increment in fluid content[3] and total stress respectively and are given by

$$\zeta = \frac{1}{M} p + \alpha \operatorname{div} \mathbf{u} \equiv \phi(\operatorname{div} \mathbf{u} - \operatorname{div} \mathbf{u}_F) \tag{A.11}$$

$$\boldsymbol{\sigma} = 2G\mathbf{E}^L + \lambda \left(\operatorname{div} \mathbf{u}\right) \mathbf{I} - \alpha p \mathbf{I} \tag{A.12}$$

where ϕ is the porosity, $\mathbf{u}_F$ is the fluid displacement and p is the pore fluid pressure. The variables M and α would be refered to as the Biot modulus and the Biot constant respectively. The Biot constant is representative of the degree of saturation of pore fluid inside the pores with $\alpha = 1$ being the fully saturated state and $\alpha < 1$ being the partially saturated state. Following the work of Geertsma [36] and Nur and Byerlee [62], the Biot modulus and Biot constant were established as

$$M = \frac{1}{c\phi + \frac{\alpha - \phi}{K_s}} \tag{A.13}$$

$$\alpha = 1 - \frac{K}{K_s} \tag{A.14}$$

where K and K_s represent the effective and grain bulk moduli respectively. We shall see in Chapter 3 that if the Biot modulus is assumed to be time-invariant, then (A.9) corresponds to the mass conservation equation for linearized slightly compressible single phase flow. Rice and Cleary [73] reworked the Biot thery in terms of parameters that were representative of the skeletal response at $t = 0$ called the ‘undrained response’ and $t \to \infty$ called the ‘drained response’. The

---

[3] The interested reader is refered to Coussy [24] for a better understanding of the fluid mass increment.

undrained response was characterized by the fluid mass increment being zero and the drained response was characterized by the pore pressure being zero. The stress-strain relations under drained conditions are identical to the ones of non - porous media, provided they are expressed in terms of the effective stress. Skempton [79] introduced a parameter called Skempton coefficient that characterizes the instantaneous response ($t = 0$) of the porous skeleton. It is the ratio between the induced pore pressure and mean stress under a confining pressure in undrained conditions as follows

$$B = -\frac{\delta p}{\delta \sigma}|_{\xi=0} \tag{A.15}$$

It is clear from the definition (A.15) that $0 \leq B \leq 1$. Bishop [11] arrived at the following relation for the Skempton coefficient

$$B = \frac{\frac{1}{K} - \frac{1}{K_s}}{\frac{1}{K} - \frac{1}{K_s} + \phi(c - \frac{1}{K_s})}$$

The undrained parameters are obtained by setting the fluid increment to zero ($\xi = 0$). From (A.11), that results in

$$p = -\alpha\, M \operatorname{div} \mathbf{u}$$

which when substituted in (A.12) results in

$$\boldsymbol{\sigma} = 2G\mathbf{E}^L + (\lambda + \alpha^2 M) \operatorname{div} \mathbf{u}\, \mathbf{I} \tag{A.16}$$

from which the undrained bulk modulus K_u can be obtained as

$$K_u = \frac{\frac{1}{3} tr(\boldsymbol{\sigma})}{tr(\mathbf{E}^L)} = \frac{2G}{3} + \lambda + \alpha^2 M = K + \alpha^2 M \tag{A.17}$$

From (A.13), (A.14) and (A.17), the undrained bulk modulus can be written as

$$K_u = K\left(1 + \frac{\alpha^2}{c\phi K + (\alpha - \phi)(1-\alpha)}\right)$$

It is easy to see that in case of a porous solid saturated with compressible pore fluid i.e. $\alpha = 1$, $c > 0$, the elastic skeleton is compressible under undrained conditions

$$\begin{aligned} K_u &= K\left(1 + \frac{1}{c\phi K}\right) \\ tr(\mathbf{E}^L) &= \operatorname{div}\mathbf{u} = \frac{\frac{1}{3}tr(\boldsymbol{\sigma})}{K_u} \neq 0 \end{aligned} \tag{A.18}$$

and in case of a porous solid saturated with incompressible pore fluid i.e. $\alpha = 1$, $c = 0$, the elastic skeleton is incompressible under undrained conditions

$$\begin{aligned} K_u &= \infty \\ tr(\mathbf{E}^L) &= \operatorname{div}\mathbf{u} = \frac{\frac{1}{3}tr(\boldsymbol{\sigma})}{K_u} = 0 \end{aligned} \tag{A.19}$$

Further, from (A.16), it is easy to see that

$$\sigma_{ij} = 2GE^L_{ij} \qquad (\text{if } i \neq j)$$

which is the same with the case of drained deformation. This means that shear modulus is the same in drained and undrained deformations.

A.1 From Navier-Stokes equations to Darcy's law

Although the Darcy's law is the most commonly used equation for flow in porous media, fluid mechanics theory suggests that fluid flow is in fact governed by the Navier-Stokes equations of balance of momentum of the fluid. The key component in arriving at a connection between the Darcy's law and the Navier-Stokes equations is the 'averaging theorem'. The theorem is based upon the well-known Reynolds transport theorem and relates the volume average of the spatial derivative to the spatial derivative of the volume average of any quantity: scalar, vector or tensor. The theorem was first presented in Slattery [81] and later elucidated upon by Whitaker [92]. Gray and Neill [40] used the averaging technique to show that the Navier-Stokes equations for flow in a deforming anisotropic porous medium reduced to the Darcy's law in the limit of slow flow.

A.2 Finite element modeling of consolidation in porous media

The first successful attempt in applying the finite element method to consolidation problem was reportedly made by Sandhu and Wilson [75] for the case of elastic porous solid saturated with incompressible pore fluid. Ghaboussi and Wilson [37] reworked the approach with the assumption of incompressibility of pore fluid and solid grains being relaxed. They used four-noded quadrilateral interpolation functions for both the pore pressure and the displacement variables but with additional incompatible interpolation functions[4]

---

[4] Wilson et al. [94] presented the idea of 'incompatible mode elements' to resolve the problem of shear locking in thin beams. Shear locking refers to the large, unphysical shear strains that arise in thin beams due to the use of standard four-noded linear quadrilateral plane stress elements.

for the latter. Booker and Small [12] established limits on the value of the time-marching coefficient for the unconditional stability of the fully discrete formulation of the equations of consolidation. Sandhu et al. [76] studied the response of a fully discrete formulation of the equations of two-dimensional consolidation of an elastic porous solid saturated with incompressible pore fluid with six-noded triangular elements and eight-noded quadrilateral elements. They reported that schemes that used equal order interpolation for pore pressure and displacement gave initial errors in pore pressure which do not dissipate in time and are oscillatory in space. Vermeer and Verruijt [88] established lower bounds on the time step size below which the fully discrete formulation yields an inaccurate pore pressure distribution with violent oscillations.

A.2.1 The Stokes' problem under undrained conditions with incompressible fluid and solid constituents

Although the Stokes' problem is not dealt with in this dissertation needlessly, this module provides a perspective on the advances in the finite element modeling of the same. From (A.4) and (A.19), it is clear that the elasticity problem under undrained conditions for the case $\alpha = 1$, $c = 0$, along with appropriate boundary conditions on $\mathbf{u}$ and p is in fact the Stokes' problem given by

$$\begin{aligned} &G \operatorname{div} \operatorname{grad} \mathbf{u} - \operatorname{grad} p = \mathbf{0} \\ &\operatorname{div} \mathbf{u} = 0 \end{aligned}$$

The work of Sandhu et al. [76] clearly suggested that the use of equal order finite element interpolation for p and $\mathbf{u}$ led to oscillatory solutions for p to the Stokes' problem. The reason lied in the failure to satisfy the well-known Ladyženskaja-Babuška-Brezzi (LBB) condition (see Brezzi and Bathe [14]) which poses a constraint on the finite element spaces used for p and $\mathbf{u}$ for the solution to be stable. Hughes et al. [49] spawned a class of methods called 'stabilized methods' meant to circumvent the deficiencies of equal order interpolation for p and $\mathbf{u}$ in obtaining a stable solution to the problem. Zienkiewicz et al. [99] studied the problem in the context of a saddle-point systems[5] and proposed an extension to the well-known patch test[6] as an alternative to the mathematically rigorous LBB condition. Murad and Loula [61] suggested the use of LBB stable Taylor-Hood elements (see Taylor and Hood [83]) and proposed a post-processing technique that improves the order of convergence of pore pressure by one. Korsawe et al. [54] reformulated the problem as a least-squares minimization problem with lowest order Taylor-Hood spaces for pressure and displacement variables and compatible Raviart-Thomas spaces[7] for fluid velocity and stress variables. Phillips and Wheeler [69] presented a formulation that combats the nonphysical pressure oscillations

---

[5] The interested reader is refered to Benzi et al. [7] for an understanding of saddle-point problems.

[6] The patch test for finite elements was first presented by the British engineer Bruce Irons at a conference on 'Matrix Methods in Structural Mechanics' held at Wright-Patterson Air Force Base, Ohio, USA on 26-28 October, 1965. The report can be found at `http://contrails.iit.edu/items/show/8575`.

[7] The interested reader is refered to Chapter III of Brezzi and Fortin [15] for a description of various mixed finite element spaces. The reader shall also find an excellent rendition of the treatment of the Stokes' problem in Chapter VI.

using Raviart-Thomas-Nedelec spaces[8] for the flow variables and a discontinuous Galerkin space[9] for the displacement variable. Liu et al. [55] presented a formulation that addresses the problem of oscillatory pore pressures using discontinuous Galerkin elements on areas with high pressure gradients and continuous Galerkin elements on the remaining domain.

A.2.2 Mixed formulation for flow with compressible fluid constituent

We have seen in the previous module that LBB stable mixed finite element spaces are a popular choice for displacement and pressure in the Stokes' equations. But, we have seen in (A.18) that if the pore fluid compressibility is strictly positive, the incompressibility constraint under undrained conditions is not operative and hence the Stokes' problem for displacement and pressure does not arise. Now, we rewrite the equation of mass conservation (A.9) for linearized slightly compressible single phase flow along with the Darcy law (A.10) for steady state conditions ($\frac{\partial \xi}{\partial t} = 0$) in the presence of gravity as

$$\begin{aligned} &\operatorname{div} \mathbf{w}^F = 0 \\ &k^{F^{-1}} \mathbf{w}^F + \operatorname{grad} p = \rho^F \mathbf{g} \end{aligned}$$

where ρ^F is the fluid density and $\mathbf{g}$ is the gravity vector. The above system represents a saddle point problem[10] for $\mathbf{w}^F$ and p. Thus, although we can

---

[8] Raviart-Thomas-Nedelec spaces are the three-dimensional analogues of the Raviart-Thomas spaces on rectangles.

[9] The interested reader is refered to Girault et al. [39] for an understanding of Discontinuous Galerkin methods.

[10] See chapter 3 of Wan [90]

avoid the necessity of the use of mixed finite element spaces for displacement and pressure by maintaining a strictly positive pore fluid compressibility, the above argument suggests a need, instead, for mixed finite element spaces for the fluid velocity and pressure under steady state conditions in the presence of gravity.

A.3 A review of solution algorithms

The early works of Sandhu and Wilson [75] and Ghaboussi and Wilson [37] along with later renditions of Jha and Juanes [51], Phillips and Wheeler [68] and Ferronato et al. [32] are refered to as fully coupled schemes where the pressures, fluid velocities and displacements are solved simultaneously. The fully coupled approach, although unconditionally stable, requires careful implementation with substantial local memory requirements and specialized linear solvers. On the other hand, the works of Park [67], Zienkiewicz et al. [100], Wheeler and Gai [91] and Kim et al. [53] are refered to as staggered solution schemes[11] in which an operator splitting strategy is used to split the coupled problem into well-posed flow and mechanics subproblems which are then solved sequentially. Carefully crafted splitting strategies would lend

---

[11] In the mid 1970s, the efforts of Ted Belytschko at Northwestern University, Thomas Hughes at California Institute of Technology and Carlos Felippa at Lockheed Palo Alto Research Laboratories would spawn the class of partitioned solution algorithms to coupled dynamical problems designed to take advantage of increasing modularity in commercial finite element codes. An excellent review of these partitioned solution procedures is given in Felippa et al. [31]. The interested reader is also refered to Yanenko and Holt [98] for an understanding of the genesis of the idea of fractional steps.

solution accuracies on par with the fully coupled approaches. An explanation of the intimate link between a fully coupled scheme and staggered solution schemes using preconditioners is presented in White et al. [93] and Castelletto et al. [21].

Previous attempts at incorporating non-matching grids for flow and mechanics include the works of Gai et al. [35], Florez et al. [34] and Castelletto et al. [22]. Gai et al. [35] reformulated a staggered solution algorithm as a special case of a fully coupled scheme and implemented the algorithm on overlapping nonmatching rectilinear grids, but avoided three dimensional intersection calculations, instead evaluating the displacement-pressure coupling submatrices using a midpoint integration rule. Florez et al. [34] implemented a procedure in which a saddle-point system with mortar spaces on nonmatching interfaces of a decomposed geomechanics domain is solved by applying a balancing Neumann-Neumann preconditioner. The procedure involves subdomain to mortar and mortar to subdomain projections, Lagrange multiplier solve and computationally expensive parallel subdomain solves at each time step. Castelletto et al. [22] implemented a procedure for nested grids in which coarse scale basis functions for the poromechanical solve are obtained in terms of fine scale basis functions by solving local equilibrium problems on each coarse scale poromechanical element. These coarse scale basis functions are then used to construct prolongation and restriction operators, which are then employed to construct a two-stage preconditioner for the coarse scale poromechanical solve.

Bibliography

[1] Y. Abousleiman, A. H. D. Cheng, L. Cui, E. Detournay, and J. C. Roegiers. Mandel's problem revisited. *Géotechnique*, 46(2):187–195, 1996.

[2] R. A. Adams. *Sobolev spaces.* Academic Press, 1975.

[3] T. Almani, K. Kumar, and M. F. Wheeler. Convergence analysis of single rate and multirate fixed stress split iterative coupling schemes in heterogeneous poroelastic media. *ICES REPORT 17-23, The University of Texas at Austin*, 2017.

[4] T. Aoki, C. P. Tan, and W. E. Bamford. Effects of deformation and strength anisotropy on borehole failures in saturated shales. *International Journal of Rock Mechanics and Mining Sciences and Geomechanics Abstracts*, 30(7):1031–1034, 1993.

[5] C. L. Bajaj, C. M. Hoffmann, R. E. Lynch, and J. E. H. Hopcroft. Tracing surface intersections. *Computer Aided Geometric Design*, 5(4): 285–307, 1988.

[6] R. Becker, P. Hansbo, and R. Stenberg. A finite element method for domain decomposition with non-matching grids. *ESAIM Mathematical Modelling and Numerical Analysis*, 37(2):209–225, 2003.

[7] M. Benzi, G. H. Golub, and J. Liesen. Numerical solution of saddle point problems. *Acta Numerica*, 14:1–137, 2005.

[8] M. A. Biot. General theory of three dimensional consolidation. *Journal of Applied Physics*, 12:155–164, 1941.

[9] M. A. Biot. Theory of elasticity and consolidation for a porous anisotropic solid. *Journal of Applied Physics*, 26(2):182–185, 1955.

[10] M. A. Biot and D. G. Willis. The elastic coefficients of the theory of consolidation. *Journal of Applied Mechanics*, 24:594–601, 1957.

[11] A. W. Bishop. The influence of an undrained change in stress on the pore pressure in porous media of low compressibility. *Géotechnique*, 23 (3):435–442, 1973.

[12] J. R. Booker and J. C. Small. An investigation of the stability of numerical solutions of biot's equations of consolidation. *International Journal of Solids and Structures*, 11(7-8):907–917, 1975.

[13] A. P. Boresi, K. P. Chong, and J. D. Lee. *Elasticity in Engineering Mechanics*. John Wiley and Sons, 3rd edition, 2010.

[14] F. Brezzi and K. J. Bathe. A discourse on the stability conditions for mixed finite element formulations. *Computer Methods in Applied Mechanics and Engineering*, 82(1-3):27–57, 1990.

[15] F. Brezzi and M. Fortin. *Mixed and Hybrid Finite Element Methods*. Springer Series in Computational Mathematics 15. Springer-Verlag New York, 1 edition, 1991.

[16] F. Brezzi, J. Douglas, R. Durán, and M. Fortin. Mixed finite elements for second order elliptic problems in three variables. *Numerische Mathematik*, 51(2):237–250, 1987.

[17] R. J. S. Brown and J. Korringa. On the dependence of the elastic properties of a porous rock on the compressibility of the pore fluid. *Geophysics*, 40(4):608–616, 1975.

[18] M. M. Carroll. An effective stress law for anisotropic elastic deformation. *Journal of Geophysical Research*, 84(B13), 1979.

[19] M. M. Carroll and N. Katsube. The role of terzaghi effective stress in linearly elastic deformation. *Journal of Energy Resources Technology*, 105(4):509–511, 1983.

[20] N. Castelletto, J. A. White, and H. A. Tchelepi. Accuracy and convergence properties of the fixed-stress iterative solution of two-way coupled poromechanics. *International Journal for Numerical and Analytical Methods in Geomechanics*, 39(14):1593–1618, 2015.

[21] N. Castelletto, J. A. White, and M. Ferronato. Scalable algorithms for three-field mixed finite element coupled poromechanics. *Journal of Computational Physics*, 327, 2016.

[22] N. Castelletto, H. Hajibeygi, and H. A. Tchelepi. Multiscale finite-element method for linear elastic geomechanics. *Journal of Computational Physics*, 331:337–356, 2017.

[23] A. H. D. Cheng. Material coefficients of anisotropic poroelasticity. *International Journal of Rock Mechanics and Mining Sciences*, 34:199–205, 1997.

[24] O. Coussy. *Poromechanics*. Wiley, 2nd edition, 2004.

[25] C. W. Cryer. A comparison of the three-dimensional consolidation theories of biot and terzaghi. *The Quarterly Journal of Mechanics and Applied Mathematics*, 16(4):401–412, 1963.

[26] L. Cui, A. H. D. Cheng, V. N. Kaliakin, Y. Abousleiman, and J. C. Roegiers. Finite element analyses of anisotropic poroelasticity: A generalized mandel's problem and an inclined borehole problem. *International Journal for Numerical and Analytical Methods in Geomechanics*, 20(6): 381–401, 1996.

[27] H. Darcy. Les fontaines publiques de la ville dijon. *Dalmont, Paris*, 1856.

[28] E. Detournay and A. H. D. Cheng. Fundamentals of poroelasticity. In *Comprehensive Rock Engineering: Principles, Practice and Projects*, volume 2, pages 113–171. Pergamon Press, 1993.

[29] N. El-Abbasi and K. J. Bathe. Stability and patch test performance of contact discretizations and a new solution algorithm. *Computers and Structures*, 79(16):1473–1486, 2001.

[30] C. A. Felippa. Iterative procedures for improving penalty function solutions of algebraic systems. *International Journal for Numerical Methods in Engineering*, 12(5):821–836, 1978.

[31] C. A. Felippa, K. C. Park, and C. Farhat. Partitioned analysis of coupled mechanical systems. *Computer methods in applied mechanics and Engineering*, 190:3247–3270, 2001.

[32] M. Ferronato, N. Castelletto, and G. Gambolati. A fully coupled 3-d mixed finite element model of biot consolidation. *Journal of Computational Physics*, 229(12):4813–4830, 2010.

[33] P. Fillunger. Österreichische wochenschrift fur den öffentlichen baudienst. *H. Lorenz: Lehrbuch der Technischen Physik, Verlag R. Oldenbourg, München und Berlin.*, 1913.

[34] H. Florez, M. F. Wheeler, A. A. Rodriguez, M. Palomino, and E. Jorge. Domain decomposition methods applied to coupled flow-geomechanics reservoir simulation. In *SPE Reservoir Simulation Symposium - The Woodlands, Texas, USA*, 2011.

[35] X. Gai, S. Sun, M. F. Wheeler, and H. Klie. A timestepping scheme for coupled reservoir flow and geomechanics on nonmatching grids. In *SPE*

Annual Technical Conference and Exhibition - (2005.10.9-2005.10.12), 2005.

[36] J. Geertsma. The effect of fluid pressure decline on volumetric changes of porous rocks. *SPE*, 210:331–340, 1957.

[37] J. Ghaboussi and E. L. Wilson. Flow of compressible fluid in porous elastic media. *International Journal for Numerical Methods in Engineering*, 5(3):419–442, 1973.

[38] R. E. Gibson, K. Knight, and P. W. Taylor. A critical experiment to examine theories of three dimensional consolidation. In *Eur. Conf. Soil Mech. Wiesbaden*, volume 1, pages 69–76, 1963.

[39] V. Girault, M. F. Wheeler, R. Glowinski, and P. Neittaanmäki. Discontinuous galerkin methods. In *Partial Differential Equations: Modeling and Numerical Simulation*, Computational Methods in Applied Sciences 16. Springer Netherlands, 2008.

[40] W. G. Gray and K. O' Neill. On the general equations for flow in porous media and their reduction to darcy's law. *Water Resources Research*, 12 (2):148–154, 1976.

[41] M. E. Gurtin, E. Fried, and L. Anand. *The Mechanics and Thermodynamics of Continua.* Cambridge University Press, 1 edition, 2010.

[42] J. O. Hallquist, G. L. Goudreau, and D. J. Benson. Sliding interfaces with contact-impact in large-scale lagrangian computations. *Computer Methods in Applied Mechanics and Engineering*, 51:107–137, 1985.

[43] Z. Hashin. Analysis of composite materials–a survey. *Journal of Applied Mechanics*, 50(3):481–505, 1983.

[44] Z. Hashin and S. Shtrikman. On some variational principles in anisotropic and nonhomogeneous elasticity. *Journal of the Mechanics and Physics of Solids*, 10(4):335–342, 1962.

[45] R. Hill. Elastic properties of reinforced solids: Some theoretical principles. *Journal of the Mechanics and Physics of Solids*, 11(5):357–372, 1963.

[46] R. Hill. A self-consistent mechanics of composite materials. *Journal of the Mechanics and Physics of Solids*, 13(4):213–222, 1965.

[47] R. Hill. On constitutive macro-variables for heterogeneous solids at finite strain. *Proceedings Mathematical Physical and Engineering Sciences*, 326 (1565):131–147, 1972.

[48] T. J. R. Hughes. *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis.* Dover Civil and Mechanical Engineering. Dover Publications, 2000.

[49] T. J. R. Hughes, L. P. Franca, and M. Balestra. A new finite element formulation for computational fluid dynamics: V. circumventing

the babuka-brezzi condition: a stable petrov-galerkin formulation of the stokes problem accommodating equal-order interpolations. *Computer Methods in Applied Mechanics and Engineering*, 59(1):85–99, 1986.

[50] R. Ingram, M. F. Wheeler, and I. Yotov. A multipoint flux mixed finite element method on hexahedra. *SIAM Journal of Numerical Analysis*, 48 (4):1281–1312, 2010.

[51] B. Jha and R. Juanes. A locally conservative finite element framework for the simulation of coupled flow and reservoir geomechanics. *Acta Geotechnica*, 2(3):139–153, 2007.

[52] N. Katsube. The constitutive theory for fluid-filled porous materials. *Journal of Applied Mechanics*, 52(1):185–189, 1985.

[53] J. Kim, H. A. Tchelepi, and R. Juanes. Stability, accuracy and efficiency of sequential methods for coupled flow and geomechanics. *SPE Journal*, 16(2):249–262, 2011.

[54] J. Korsawe, G. Starke, W. Wang, and O. Kolditz. Finite element analysis of poro-elastic consolidation in porous media: Standard and mixed approaches. *Computer Methods in Applied Mechanics and Engineering*, 195(9-12):1096–1115, 2006.

[55] R. Liu, M. F. Wheeler, C. N. Dawson, and R. H. Dean. On a coupled discontinuous/continuous galerkin framework and an adaptive penalty

scheme for poroelasticity problems. *Computer Methods in Applied Mechanics and Engineering*, 198(41-44):3499–3510, 2009.

[56] J. Mandel. Consolidation des sols (étude mathématique)*. *Géotechnique*, 3(7):287–299, 1953.

[57] T. W. McDevitt and T. A. Laursen. A mortar-finite element formulation for frictional contact problems. *International Journal for Numerical Methods in Engineering*, 48(10):1525–1547, 2000.

[58] A. Mikelić and M. F. Wheeler. Convergence of iterative coupling for coupled flow and geomechanics. *Computational Geosciences*, 17(3):455–461, 2013.

[59] A. Mikelić, B. Wang, and M. F. Wheeler. Numerical convergence study of iterative coupling for coupled flow and geomechanics. In *13th European conference on the Mathematics of Oil recovery, Biarritz, France*, 2012.

[60] A. Mikelić, B. Wang, and M. F. Wheeler. Numerical convergence study of iterative coupling for coupled flow and geomechanics. *Computational Geosciences*, 18(3):325–341, 2014.

[61] M. A. Murad and A. F. D. Loula. On stability and convergence of finite element approximations of biot's consolidation problem. *International Journal for Numerical Methods in Engineering*, 37(4):645–667, 1994.

[62] A. Nur and J. D. Byerlee. An exact effective stress law for elastic deformation of rock with fluids. *Journal of Geophysical Research*, 76(26): 6414–6419, 1971.

[63] J. T. Oden and L. Demkowicz. *Applied Functional Analysis*. Chapman and Hall, 2nd edition, 2010.

[64] P. Papadopoulos and R. L. Taylor. A mixed formulation for the finite element solution of contact problems. *Computer Methods in Applied Mechanics and Engineering*, 94(3):373–389, 1992.

[65] P. Papadopoulos and R. L. Taylor. A simple algorithm for three-dimensional finite element analysis of contact problems. *Computers and Structures*, 46(6):1107–1118, 1993.

[66] H. Parisch. A consistent tangent stiffness matrix for three-dimensional non-linear contact analysis. *International Journal for Numerical Methods in Engineering*, 28(8):1803–1812, 1989.

[67] K. C. Park. Stabilization of partitioned solution procedure for pore fluid-soil interaction analysis. *International Journal for Numerical Methods in Engineering*, 19(11):1669–1673, 1983.

[68] P. J. Phillips and M. F. Wheeler. A coupling of mixed and continuous galerkin finite element methods for poroelasticity ii: the discrete-in-time case. *Computational Geosciences*, 11(2):145–158, 2007.

[69] P. J. Phillips and M. F. Wheeler. A coupling of mixed and discontinuous galerkin finite-element methods for poroelasticity. *Computational Geosciences*, 12(4):417–435, 2008.

[70] M. J. D. Powell. Algorithms for nonlinear constraints that use lagrangian functions. *Mathematical Programming*, 14(1):224–248, 1978.

[71] M. A. Puso and T. A. Laursen. Mesh tying on curved interfaces in 3d. *Engineering Computations*, 20(3):305–319, 2003.

[72] M. A. Puso and T. A. Laursen. A mortar segment-to-segment contact method for large deformation solid mechanics. *Computer Methods in Applied Mechanics and Engineering*, 193(6-8):601–629, 2004.

[73] J. R. Rice and M. P. Cleary. Some basic stress diffusion solutions for fluid-saturated elastic porous media with compressible constituents. *Reviews of Geophysics*, 14(2):227–241, 1976.

[74] S. Saeb, P. Steinmann, and A. Javili. Aspects of computational homogenization at finite deformations: A unifying review from reuss' to voigt's bound. *Applied Mechanics Reviews*, 68(5):050801–1–050801–33, 2016.

[75] R. S. Sandhu and L. Wilson. Finite-element analysis of seepage in elastic media. *Journal of the Engineering Mechanics Division*, 95(3):641–652, 1969.

[76] R. S. Sandhu, H. Liu, and K. J. Singh. Numerical performance of some finite element schemes for analysis of seepage in porous elastic media.

International Journal for Numerical and Analytical Methods in Geomechanics, 1(2):177–194, 1977.

[77] H. Si. Tetgen, a delaunay-based quality tetrahedral mesh generator. *ACM Trans. on Mathematical Software*, 41(2), 2015.

[78] J. C. Simo, P. Wriggers, and R. L. Taylor. A perturbed lagrangian formulation for the finite element solution of contact problems. *Computer Methods in Applied Mechanics and Engineering*, 50(2):163–180, 1985.

[79] A. W. Skempton. The pore-pressure coefficients a and b. *Géotechnique*, 4(4):143–147, 1954.

[80] A. W. Skempton. Significance of terzaghi's concept of effective stress (terzaghi's discovery of effective stress). In *From theory to practice in soil mechanics (L. Bjerrum and A. Casagrande and R. B. Peek and A. W. Skempton, eds.). New York-London: John Wiley and Sons 1960*, 1960.

[81] J. C. Slattery. Flow of viscoelastic fluids through porous media. *AIChE Journal*, 13(6):1066–1071, 1967.

[82] J. M. Steele. *The Cauchy-Schwarz Master Class: An Introduction to the Art of Mathematical Inequalities*. Maa Problem Books Series. Cambridge University Press, 2004.

[83] C. Taylor and P. Hood. A numerical solution of the navier-stokes equations using the finite element technique. *Computers and Fluids*, 1(1): 73–100, 1973.

[84] K. Terzaghi. Die berechnung der durchl ssigkeitsziffer des tones aus dem verlauf der hydrodynamischen spannungserscheinungen. *Sitzungsber. Akad. Wiss. (Wien), Math.-Naturwiss*, 132:125–138, 1923.

[85] K. Terzaghi. The shearing resistance of saturated soils and the angle between the planes of shear. *First Int. Conf. Soil Mech., Vol. 1, Harvard University*, pages 54–56, 1936.

[86] M. Thompson and J. R. Willis. A reformation of the equations of anisotropic poroelasticity. *Journal of Applied Mechanics*, 58(3):612–616, 1991.

[87] T. C. T. Ting. *Anisotropic elasticity: Theory and applications.* The Oxford Engineering Science Series. Oxford University Press, 1996.

[88] P. A. Vermeer and A. Verruijt. An accuracy condition for consolidation by finite elements. *International Journal for Numerical and Analytical Methods in Geomechanics*, 5(1):1–14, 1981.

[89] A. Verruijt. Elastic storage of aquifers. In *Flow through porous media (edited by R. J. M. DeWiest) New York: Academic Press*, pages 331–376, 1965.

[90] J. Wan. *Stabilized finite element methods for coupled geomechanics and multiphase flow*. PhD thesis, Stanford University, 2002.

[91] M. F. Wheeler and X. Gai. Iteratively coupled mixed and galerkin finite element methods for poro-elasticity. *Numerical Methods for Partial Differential Equations*, 23(4):785–797, 2007.

[92] S. Whitaker. Advances in theory of fluid motion in porous media. *Industrial and Engineering Chemistry*, 61(12):14–28, 1969.

[93] J. A. White, N. Castelletto, and H. A. Tchelepi. Block-partitioned solvers for coupled poromechanics: A unified framework. *Computer Methods in Applied Mechanics and Engineering*, 303:55–74, 2016.

[94] E. L. Wilson, R. L. Taylor, W. P. Doherty, and J. Ghaboussi. Incompatible displacement models. In *Numerical and Computer Methods in Structural Mechanics, New York: Academic Press*, pages 43–57, 1973.

[95] P. Wriggers. *Computational Contact Mechanics*. Springer, 2nd edition, 2006.

[96] P. Wriggers and J. C. Simo. A note on tangent stiffness for fully nonlinear contact problems. *International Journal for Numerical Methods in Biomedical Engineering*, 1(5):199–203, 1985.

[97] P. Wriggers and G. Zavarise. A formulation for frictionless contact problems using a weak form introduced by nitsche. *Computational Mechanics*, 41(3):407–420, 2008.

[98] N. N. Yanenko and M. Holt. *The Method of Fractional Steps: The Solution of Problems of Mathematical Physics in Several Variables*. Springer-Verlag Berlin Heidelberg, 1 edition, 1971.

[99] O. C. Zienkiewicz, S. Qu, R. L. Taylor, and S. Nakazawa. The patch test for mixed formulations. *International Journal for Numerical Methods in Engineering*, 23(10):1873–1883, 1986.

[100] O. C. Zienkiewicz, D. K. Paul, and A. H. C. Chan. Unconditionally stable staggered solution procedure for soil-pore fluid interaction problems. *International Journal for Numerical Methods in Engineering*, 26 (5):1039–1055, 1988.

[101] O. C. Zienkiewicz, R. L. Taylor, and J. Z. Zhu. *The Finite Element Method: Its Basis and Fundamentals*. Butterworth-Heinemann, 7th edition, 2013.

[102] T. I. Zohdi and P. Wriggers. *Introduction to computational micromechanics*. Lecture notes in applied and computational mechanics, v. 20. 2005.

Vita

Saumik Dana was born in Surat, Gujarat, India on January 6, 1987 and moved to Mumbai, Maharashtra, India with family on October 1987. He graduated with a Bachelor of Engineering degree in Mechanical Engineering from University of Mumbai in 2008. He obtained a Master of Engineering degree in Mechanical Engineering from Indian Institute of Science, Bangalore in 2011. He continued as a Junior Research Fellow at the same institute for a year. He joined graduate school at The University of Texas at Austin in Fall 2012 and obtained a Master of Science degree in Engineering Mechanics in Fall 2014. He started his affiliation with the Center for Subsurface Modeling at the Institute for Computational Engineering and Sciences in pursuit of a Doctor of Philosophy degree in Engineering Mechanics in Spring 2015.

Permanent address: saumik@utexas.edu

This dissertation was typeset with LaTeX[†] by the author.

[†] LaTeX is a document preparation system developed by Leslie Lamport as a special version of Donald Knuth's TeX Program.